



# **Risco de Crédito**

por

Vítor Fernando da Silva Gonçalves Ferreira

Tese de Mestrado em Economia

Orientado por: Professor Doutor José Manuel Peres Jorge

2014

## **Nota Bibliográfica**

Vítor Fernando da Silva Gonçalves Ferreira.

Nasci a 21 de Outubro de 1983 em Celorico de Basto. Sou casado, pai de um filho e terminei a licenciatura em Economia na Universidade de Trás-os-Montes e Alto Douro em 2006. No mesmo ano iniciei a minha atividade profissional num Banco Privado em Portugal, situação que se mantém até ao momento.

No ano letivo 2011-2012 iniciei a parte curricular do Mestrado em Economia da qual faço agora a tese.

## Resumo

Considerando o impacto que o financiamento às empresas proporciona no crescimento da economia, surge de imediato que critérios terá o setor financeiro para proceder a esse financiamento. O desejável para o setor financeiro é as empresas apresentarem bons indicadores financeiros mas a seleção desses indicadores pode ser feita com recurso a técnicas estatísticas diferenciadas.

O objetivo deste trabalho é selecionar um conjunto de indicadores financeiros capazes de sinalizar sinais de fraqueza económica nas empresas e assim contribuir como uma ferramenta disponível na análise de concessão de crédito. Esta seleção é efetuada com recurso à análise fatorial e à regressão logística.

A análise fatorial e logística foram utilizadas, por exemplo, por West em 1984 para prever precocemente sinais de fraquezas nos balanços de bancos comerciais.

Os modelos sobre previsão de falência e a parametrização do crédito surgiram na segunda metade do século XX. Ainda hoje alguns métodos de previsão de falência desenvolvidos na década de 70 têm uma aplicabilidade muito fiável, veja-se, por exemplo o modelo *ZETA* desenvolvido por Altman e co-autores (1977).

Outros modelos têm sido desenvolvidos por grandes instituições financeiras e de Rating na gestão de risco interno associado a grandes carteiras de crédito. São disso exemplos os modelos *CreditRisk* e o *CreditMetrics*.

Este trabalho pretende contribuir para a seleção de indicadores que melhor sinalizem sinais de fraqueza nos balanços das pequenas e médias empresas em Portugal com base numa amostra de 500 empresas industriais nacionais.

Numa primeira abordagem os fatores “Encargos e Responsabilidades” e os “Proveitos Operacionais Correntes” são os que melhor predizem a situação financeira das empresas.

Códigos JEL: C25, G21, G33

Palavras-chave: *Credit Risk*, *Credit Scoring* e Análise de Crédito.

## Abstract

Attending to the impact of the enterprises' financing in economic growth, it is important to know which criteria has the financial sector to do it. Good financial indicators of the enterprises are the criteria and their selection can be done by different statistical techniques.

The objective of this work is to select some financial indicators of economic weakness of the enterprises to be available in the credit grant. This selection is done by factorial analysis and logistic regression, also used by West (1984) to predict weakness signals in the balance of commercial banks.

The models about bankruptcy prediction and credit parameterization raised in the second half of twentieth century and in the seventies are still reliable, like for example the *ZETA* model developed by Altman and co-workers (1977).

Other models have been developed by big financial and rating institutions in the management of internal risk associated to large credit portfolio, such as “CreditRisk” and “CreditMetrics”.

This work will contribute to select the indicators that better indicate weakness signals in the balance sheet of small and medium enterprises in Portugal based in a sample of 500 national industrial enterprises.

The factors “Charges and Responsibilities” and “Current Operating Income” seems to be the best ones to predict the financial situation of the enterprises.

Codes: C25, G21, G33

Key-words: *Credit Risk*, *Credit Scoring* and Credit analysis.

## **Agradecimentos**

Ao meu filho Tiago e à minha mulher Odete

Ao Professor Doutor José Manuel Peres Jorge

## **Siglas**

AF	Autonomia Financeira
CAE	Código de Atividade Empresarial
CMVMC	Custo da Mercadoria Vendida e Matéria Consumida
CP	Curto Prazo
CPessoal	Custo com o Pessoal
EBE	Excedente Bruto de Exploração
EBIT	Earnings Before Interest and Taxes
EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization
FSE	Fornecimento e Serviços Externos
DP e Caixa	Depósitos Bancários e Caixa
GAC	Grau de Alavancagem Combinada
GAF	Grau de Alavancagem Financeira
GAO	Grau de Alavancagem Operacional
IMP	Impostos
MLP	Medio Longo Prazo
N Empregados	Número de Empregados
NFM	Necessidade de Fundo de maneo
OLS	Ordinary Least Squares
PMP	Prazo Médio de Pagamento
PME's	Pequenas e Médias Empresas
PMR	Prazo Médio de Recebimento
Passivo-CP	Passivo de Curto Prazo
Passivo-MLP	Passivo de Médio Longo Prazo
PSU	Proveitos Suplementares
Rendibilidade-CP	Rendibilidade Curto Prazo
RL	Resultado Líquido
RLE	Resultado Líquido do Exercício
RO	Resultado Operacional
RAI	Resultado Antes de Impostos
VAB	Valor Acrescentado Bruto
VN	Volume de Negócios

## Índice

1- Introdução .....	1
2- Enquadramento económico .....	6
3- Revisão de Literatura.....	9
4- Descrição das técnicas estatísticas mais utilizadas na literatura .....	14
4.1- Análise Discriminante .....	14
4.2- Modelo Logit .....	16
5- Descrição da amostra.....	17
6- Definição das variáveis.....	18
6.1- Estatísticas descritivas das variáveis .....	21
7- Análise em componentes principais: a importância no tratamento de dados .....	23
7.1- Aplicação da ACP na amostra: distribuição de variáveis e indivíduos .....	24
8- Análise fatorial .....	28
8.1- Descrição da técnica .....	28
8.2- Aplicação na amostra.....	29
8.3- Pressupostos iniciais .....	32
8.4- Resultados obtidos com a análise fatorial.....	32
9- Modelos de Regressão .....	42
9.1- O modelo Logit e a variável dependente “Situação” .....	43
9.2- O Modelo Logit e a variável dependente “Incidentes” .....	47
9.3- O Modelo OLS e a variável dependente “Endividamento” .....	51
10- Interpretação dos resultados .....	53
11- Conclusão .....	58
12- Referências bibliográficas .....	59
13- Anexos .....	61

## Índice de Tabelas

Tabela 1: Variância Total Explicada .....	33
Tabela 2: “Encargos e Responsabilidades” .....	34
Tabela 3: “Resultados Obtidos” .....	35
Tabela 4: “Solvabilidade” .....	35
Tabela 5: “Liquidez” .....	36
Tabela 6: “Proveitos Operacionais Correntes” .....	37
Tabela 7: “Rendibilidade” .....	37
Tabela 8: “Prazos de Pagamentos” .....	38
Tabela 9: “Custos Operacionais Correntes” .....	38
Tabela 10: “Volatilidade dos Resultados Operacionais” .....	39
Tabela 11: “Sensibilidade dos Resultados Correntes” .....	39
Tabela 12: “Variabilidade das Vendas” .....	40
Tabela 13: “Capacidade de Reembolso” .....	40
Tabela 14: Output do SPSS da Regressão Logística: variável dependente “Situação” .	44
Tabela 15: Output capacidade preditiva do modelo: variável dependente “Situação” ...	46
Tabela 16: Output SPSS da Regressão Logística: variável dependente “Incidentes” ....	47
Tabela 17: Output da predição do modelo: variável dependente “Incidentes” .....	50
Tabela 18: Output do Eviews regressão OLS: variável dependente “Endividamento” ..	52



## **Índice de Figuras**

Figura 1: Variáveis Seleccionadas.....	18
Figura 2: Distribuição das Variáveis .....	24
Figura 3: Distribuição dos Indivíduos .....	24
Figura 4: Regressões utilizadas neste estudo.....	42

## **Índice de Quadros**

Quadro 1: Principais Técnicas Estatísticas utilizadas ao longo dos anos .....	9
Quadro 2: Estatísticas descritivas das variáveis da amostra .....	21

## **Índice de Gráficos**

Gráfico 1: Crédito total concedido à economia .....	5
Gráfico 2: Taxa de juro aplicada a novas operações de crédito às empresas .....	6
Gráfico 3: Variabilidade crédito cobrança duvidosa .....	6
Gráfico 4: Gráfico de sedimentação .....	31
Gráfico 5: A variável “Situação” em função dos “Encargos e Responsabilidades” .....	43
Gráfico 6: A variável “Situação” em função dos “Proveitos Operacionais Correntes” .	43
Gráfico 7: A variável “Incidentes” em função dos “Prazos de Pagamentos” .....	47

## 1- Introdução

As famílias, as empresas e o Estado têm um papel muito importante na economia pois quando um destes agentes passa por dificuldades, o risco de contágio aos outros é elevado. Por exemplo, entrando em dificuldades económicas, as famílias vão racionalizar as suas decisões e, entre muitas, poderão decidir não comprar cereais para o pequeno-almoço. Consequentemente, a vendedora da mercearia deixará de vender cereais por não haver procura e decidirá não encomendar mais ao fornecedor. Este, por sua vez, deixará de produzir cereais por não haver mercado para os vender e suspenderá as encomendas de milho e trigo ao agricultor. Finalmente, o agricultor, não tendo a quem vender os cereais, deixará de cultivá-los e, consequentemente, não tendo dinheiro para pagar os salários dos seus funcionários acabará por despedi-los. Entretanto, os funcionários despedidos recorrerão à Segurança Social para usufruírem do subsídio de desemprego e a Segurança Social ficará sem contribuintes, aumentando o número de beneficiários das prestações sociais. Perante o desequilíbrio de saldos, o Estado será obrigado a suportar os mesmos. Este exemplo é radical mas permite perceber a dependência entre os 3 agentes referidos acima.

Assim, perante esta realidade, e tendo presente a sua importância na economia, não é possível isolar, individualmente, o que se passa em cada um destes agentes. As notícias que surgem nos meios de comunicação mostram que a existência de dificuldades financeiras afetam toda a sociedade. A título exemplificativo temos três notícias que espelham essa interligação entre os vários agentes económicos e a sua importância. Em primeiro, *As famílias e as empresas portuguesas estão com cada vez maiores dificuldades para honrar os seus compromissos junto da banca, com o crédito malparado a voltar a crescer em maio, de acordo com os dados do Banco de Portugal.* [DN, 2011-07-21]

Em segundo, *Nos últimos anos, o Banco de Portugal tem vindo a fazer várias inspeções às carteiras de crédito dos oito maiores bancos a operar em Portugal (...). Os setores da construção e do imobiliário, onde se registam níveis elevados de crédito malparado, têm recebido especial atenção do supervisor liderado por Carlos Costa.* [TVI 24, 2013-06-26 18:46]

Por fim, o *Banco de Portugal afirma que a carga fiscal conjugada com cortes nos salários e pensões vai afetar rendimento disponível em 2014. E avisa que a consolidação orçamental não é amiga do crescimento do emprego e constitui uma ameaça à própria estabilidade financeira do sistema bancário.* [DN, 2013-11-27]

Uma nota importante a retirar nesta relação é a importância que as empresas assumem no centro da sociedade. Estas são em primeiro lugar responsáveis pela criação de emprego mas também são responsáveis pela produção de bens e serviços para o consumidor final. Existem vários tipos de empresas, entre as quais: empresas em que o dono está inscrito nas finanças como empresário em nome individual; outras, designadas de sociedades por quotas, em que no mínimo existem dois sócios; e outras ainda, ditas sociedades anónimas, em que o capital social está distribuído por um grupo de acionistas anónimos. As empresas podem classificar-se em pequenas e médias empresas (PME's) e grandes empresas. Consideram-se como PME's aquelas que têm até 250 funcionários e um volume de faturação inferior a 50 milhões de euros. As grandes empresas, por seu lado, são aquelas em que o número de funcionários é superior a 250 e o volume de faturação ultrapassa os 50 milhões de euros. Por fim, as empresas podem apresentar dois estados: solvente (em plena atividade comercial) e insolvente (em falência ou em reestruturação/recuperação judicial). Relativamente ao setor de atividade económico das empresas, este pode ser primário (atividade mais ligada à natureza/agricultura), secundário (atividade ligada à indústria de transformação, construção e produção de energia) ou terciário (normalmente ligado ao comércio, atividades financeiras, turismo e transportes).

Perante as dificuldades apresentadas, salienta-se a importância que o setor bancário tem assumido, ao longo dos anos, como parceiro das famílias, das empresas e até do Estado. As dificuldades económicas que envolvem a sociedade afetam diretamente os bancos enquanto entidades financeiras. Assim, perante as possíveis dificuldades, o setor bancário tem vindo a proteger-se com a introdução de mecanismos cada vez mais sofisticados nos processos de decisão de crédito e de previsão de falências para salvaguardar a sua própria solvabilidade e liquidez. A questão do risco de crédito toma assim uma importância acrescida no setor financeiro e a literatura realça essa importância. Altman e Saunders (1997) referem que o cálculo do risco de crédito tem uma especial atenção no setor bancário e tem sofrido alterações muito significativas

ao longo dos últimos 20 anos. Também Lopez e Saidenberg (2000) referem que os Bancos têm investido muitos recursos no desenvolvimento de modelos de risco para assim medirem o próprio risco financeiro e o capital necessário para cobrir esse risco. Existem várias técnicas para prever a probabilidade de falências das empresas entre as quais: a análise discriminante, as árvores de decisão, as redes neurais, a programação linear, a análise fatorial, a regressão logística (Logit) e a regressão linear (OLS).

Desta forma, o principal objetivo deste trabalho é selecionar um conjunto de variáveis e indicadores financeiros que sejam uma mais-valia na análise do risco de crédito. Estes indicadores devem permitir interpretar a probabilidade de uma empresa entrar num estado insolvente, bem como contribuir na verificação da probabilidade de uma empresa registar incidentes ao nível dos pagamentos a terceiros. Por fim, esses indicadores devem ser capazes de avaliar o nível de endividamento das empresas.

Para a elaboração deste estudo recorri à base de dados SABI (Sistema de Análise de Balanços Ibérios) de onde selecionei, aleatoriamente, uma amostra de 500 empresas industriais e extraí informação referente a 44 variáveis associadas a cada empresa. O ano de referência é 2012. No tratamento dos dados recorri à análise fatorial. Esta técnica permite agrupar as variáveis de acordo com a sua natureza e a cada grupo de variáveis que se forma dá-se o nome de fator ou componente principal. Assim, cada fator representa um grupo de variáveis correlacionadas entre si. Após obter os fatores, procedi à interpretação probabilística de uma empresa se tornar insolvente bem como a probabilidade de registar incidentes ao nível dos seus pagamentos. Nesta fase, utilizei a regressão logística pelo motivo de as variáveis dependentes assumirem apenas dois valores, “Situação” (Solvente=0 e Insolvente=1) e ”Incidentes” (Sem incidentes=0 e Com incidentes=1). Como variáveis independentes utilizei os fatores obtidos na análise fatorial. Para o estudo do nível de endividamento, usei a regressão linear dos mínimos quadrados ordinários. A variável dependente é o “Endividamento” e as variáveis explicativas são os fatores obtidos na análise fatorial.

A análise fatorial e a regressão logística já foram utilizadas no passado, por exemplo por West em 1984, na construção de um sistema de alerta precoce de fraqueza para os bancos comerciais dos Estados Unidos da América. Usei a análise fatorial neste estudo que me permitiu simplificar a interpretação da informação e reagrupou 44 variáveis iniciais em apenas 12 grupos, agora designados por fatores: “Encargos e

Responsabilidades”, “Resultados Obtidos”, “Solvabilidade”, “Liquidez”, “Proveitos Operacionais Correntes”, “Rendibilidade”, “Prazos de Pagamento”, “Custos Operacionais Correntes”, “Volatilidade dos Resultados Operacionais”, “Sensibilidade dos Resultados Correntes”, “Variabilidade das Vendas” e a “Capacidade de Reembolso”. Este conjunto de fatores é responsável por 80% de toda a variância explicada e os fatores “Encargos e Responsabilidades” e os “Proveitos Operacionais Correntes” são os que melhor classificam corretamente as empresas em situação de solvência e insolvência.

O fator “Encargos e Responsabilidades” engloba quinze variáveis iniciais e é responsável por 36,10% de toda a inércia explicada. O sinal positivo do coeficiente associado a este fator indica que quanto maior forem os “Encargos e Responsabilidades” maior será a probabilidade para a ocorrência de insolvência e incidentes. No que respeita à exposição do endividamento, os “Encargos e Responsabilidades” apresentam um sinal negativo no seu coeficiente de regressão o que permite concluir que quanto maior for o valor nesta variável menor tenderá a ser o aumento do endividamento por parte da empresa. O fator “Proveitos Operacionais Correntes” é formado pelas variáveis “Proveitos Operacionais por Empregado” e “CMVMC”. Este fator contém 93% de toda a informação inicial das variáveis e contribui positivamente para ocorrência de insolvência e negativamente para a probabilidade de registo de incidentes. Ao nível do endividamento quanto maior for o fator “Proveitos Operacionais Correntes” menor será a aumento do endividamento.

O modelo Logit aplicado no estudo da variável “Situação”, e explicado pelos 12 fatores, apresenta uma taxa de acerto de 97,40%. Isto significa que das 490 empresas, o modelo classificou 3 como sendo “Insolventes” quando eram “Solventes”, dando uma margem de erro de apenas 2,60%.

No que respeita apenas à probabilidade de ocorrência de incidentes nos pagamentos, os fatores “Resultados Obtidos” e “Prazos de Pagamento” são os que apresentam um melhor desempenho. O sinal negativo do coeficiente associado ao fator “Resultados Obtidos” indica que a probabilidade de ocorrência de incidentes tende a diminuir quanto maior forem os “Resultados Obtidos” e comparativamente a uma outra empresa nas mesmas condições. O coeficiente associado à variável “Prazos de Pagamento” é positivo, o que pode significar que a probabilidade de ocorrência de

incidentes tende a ser superior numa empresa que apresente um prazo de pagamento e recebimento mais lato.

O modelo Logit utilizado no estudo da variável “Incidentes” apresenta uma taxa de acerto de 77,80% e uma taxa de erro de 22,20%. Na interpretação do endividamento, os 12 fatores associados aos seus coeficientes de regressão, e à exceção do fator “Variabilidade das Vendas”, são individualmente significativos. Ainda assim, o coeficiente associado ao fator “Variabilidade das Vendas” é negativo o que parece sinalizar que quanto maior for a variabilidade positiva nas vendas menor será o aumento do endividamento, para tudo o resto constante.

Assim, o relatório terá inicialmente uma contextualização do tema (secção 2) e uma revisão da literatura (secção 3). De seguida, serão descritas as técnicas utilizadas: análise discriminante e logística (secção 4). Posteriormente, será feita a descrição da amostra (secção 5) e definição das variáveis (secção 6). A análise em componentes principais será aplicada de seguida nos dados da amostra (secção 7). Na secção 8, proceder-se-á ao tratamento de dados com a aplicação da análise fatorial. O modelo Logit e OLS serão utilizados na secção 9 para verificar o efeito dos fatores retidos na análise fatorial como variáveis explicativas em três contextos diferentes de variáveis dependentes. Por fim, será feita uma interpretação dos resultados (secção 10) e uma conclusão (secção 12).



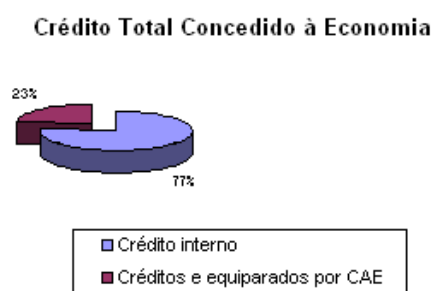
## 2- Enquadramento económico

Desde 2011 que Portugal tem vindo a fazer um ajustamento da economia dando enfoque especial na consolidação orçamental e das contas nacionais. É um período onde se assiste a mudanças ao nível laboral, na educação, na saúde, na justiça e na recapitalização dos balanços de todo o sistema financeiro. As famílias confrontam-se com uma maior carga fiscal e consequentemente perda salarial. O número de desempregados aumentou, fruto da falência de muitas empresas, e um clima de desconfiança entre os vários agentes económicos surgiu com a nova realidade económica. Fruto da recessão económica, o nível de contratação de crédito diminuiu e nem os mínimos históricos registados nas Euribor a 3, 6 e 12 meses levou a que o cenário se invertesse.

No final de 2012, a dívida total das empresas não financeiras era de 311.773 milhões de euros sendo que nas Pequenas e Médias Empresas (PME's) a dívida era de 170.969 milhões euros. Na mesma altura, os empréstimos concedidos às sociedades não financeiras foram de 106.524 milhões de euros, dos quais 79.543 milhões de euros foram atribuídos às PME's. Ao nível das imparidades, o resultado do crédito em incumprimento era de 10% nas sociedades não financeiras e de 12,4% nas PME's.

Em 2013, o setor empresarial não financeiro absorveu 23% (99.296 milhões de euros) de todo o montante de crédito concedido às empresas (330.823 milhões de euros) (Gráfico 1).

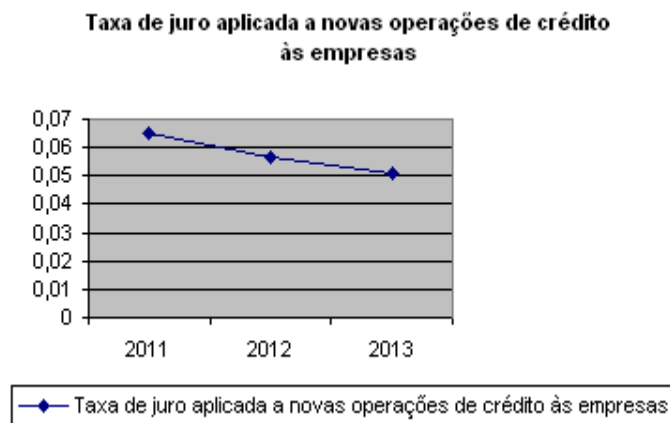
Gráfico 1- Crédito total concedido à economia.



Fonte: Banco de Portugal.

Entre 2011 e 2013, as taxas médias de juro aplicadas a novas operações de financiamento diminuíram de 6,5% para 5% (Gráfico 2) o que, no entanto, não se refletiu num aumento do montante concedido ao setor empresarial.

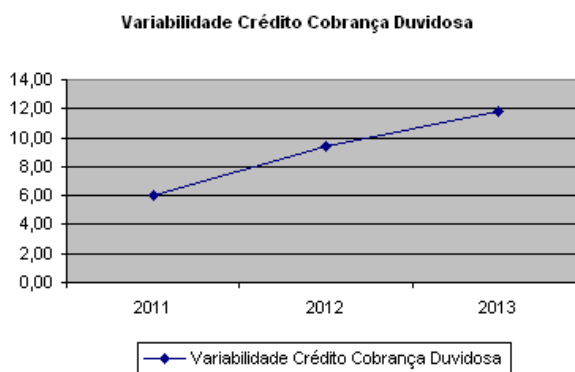
Gráfico 2- Taxa de juro aplicada a novas operações de crédito às empresas.



Fonte: Banco de Portugal.

Neste intervalo de tempo, o rácio de crédito vencido duplicou dos 6%, em 2011, para 12% em 2013 (Gráfico 3). Em termos reais, os 12 % correspondem a 11.690 milhões de euros em risco de não serem recuperados.

Gráfico 3- Variabilidade crédito cobrança duvidosa.



Fonte: Banco de Portugal.

Uma nota importante sobre a evolução da concessão de novos empréstimos e o aumento de crédito em cobrança duvidosa é que, apesar deste aumento de crédito em risco, se considerado o montante concedido, conclui-se que existe um grande controlo do rácio do crédito vencido.

A literatura vai-nos informando sobre as preocupações do setor bancário. Por exemplo, Altman e Saunders (1997) afirmam que *o setor bancário tem tido uma preocupação crescente na análise do crédito e tem investido cada vez mais em meios capazes de ajudarem na decisão de concessão de crédito. Esta metodologia é um forte contributo para a solvabilidade de um sistema financeiro.*

### 3- Revisão de Literatura

Este estudo tem associado três conceitos-chave: *Credit Risk*, *Credit Soring* e Análise de Crédito.

O *Credit Risk* define-se como a exposição que a Entidade Financiadora assume perante a concessão de um empréstimo, por exemplo, a possibilidade de incumprimento, retorno abaixo do esperado, etc. Mas também existe risco para o contraente, como uma quebra na sua atividade pode levar a empresa a dificuldades e à insolvência, entre outros.

O *Credit Scoring* permite a classificação, em pontuação, dos indivíduos, particulares ou empresas, posteriormente utilizada como suporte à tomada de decisão. Contempla o Grau de Investimento [AAA até BBB] e o Grau de Especulação [BB até D].

A **Análise de Crédito** é o processo pelo qual o decisor averigua a capacidade do contraente pagar o valor solicitado e a Entidade Financiadora de gerir essa concessão e calcular o retorno esperado.

O tema do *Credit Scoring* teve uma evolução científica muito significativa desde a década de 60 (Quadro 1).

Quadro 1: Principais Técnicas Estatísticas utilizadas ao longo dos anos.

Beaver (1966)	Análise Discriminante Univariada
Altman (1968) e Altman <i>et al.</i> (1977)	Análise Discriminante Multivariada
Martin (1977)	Análise Discriminante e o Método Logit
West (1984)	Análise Fatorial e o Método Logit
Platt e Platt (1991)	Método de Regressão Não Linear
Lawrence <i>et al.</i> (1992)	Método Logit
Smith e Lawrence (1995)	Método OLS e Logit
Lehmann (2003)	Método Logit e a informação qualitativa
Altman e Sabato(2007)	Modelo Z-Score e o Método Logit

Beaver (1966) aplicou, pela primeira vez, a técnica da análise discriminante univariada em rácios financeiros de empresas para previsão de falência. Pode dizer-se

que foi pioneiro a utilizar rácios financeiros neste tipo de análise e a partir de Beaver muitos outros estudos se focaram nos rácios financeiros.

Altman (1968) desenvolveu o modelo *Z-Score*. O modelo foi construído a partir da aplicação da análise discriminante multivariada num conjunto de cinco categorias de variáveis: liquidez, rentabilidade, endividamento, solvabilidade e índices de atividade. A amostra era constituída por 66 empresas nas quais metade insolventes e outra metade solvente. Esta técnica permite medir quais as variáveis que apresentam uma maior significância quanto à sua grandeza, situação não evidente em estudos anteriores com base na análise univariada. O *Z-Score* permite prever falências até 2 anos antes da ocorrência. Com outros autores, Altman desenvolveu o modelo *Zeta* em 1977. Este modelo é uma ampliação do anterior *Z-Score* tanto ao nível do número de empresas utilizadas na amostra, 111 empresas, como ao nível da sua natureza: industriais e de retalho. Existe também um refinamento dos dados financeiros e o número de variáveis utilizadas. A análise discriminante multivariada utilizada neste estudo permitiu prever falências até 5 anos antes. O modelo *Zeta* é ainda hoje uma referência para muitos estudos.

Martin (1977) utilizou a regressão logística e a análise discriminante na construção de um modelo de alerta precoce na falência de Bancos tendo sido pioneiro a utilizar o modelo Logit em estudos de alertas associados a Bancos. As variáveis utilizadas nas regressões são classificadas em 4 grupos: risco dos ativos, liquidez, adequação do capital e ganhos.

West (1984) explorou uma nova abordagem para a criação de um sistema de alerta precoce de fraqueza no setor bancário. No seu estudo recorreu a dois métodos: a análise fatorial e a regressão logística. Os resultados obtidos tornaram-se muito promissores para este tipo de abordagem. O autor refere que até à data só existiam duas categorias para medir o risco: os modelos de previsão de falência e os modelos de alerta de fraqueza financeira, sendo este último a sua âncora no estudo. Uma especificidade da análise fatorial é classificar grupo de variáveis no mesmo eixo principal de acordo com o nível de correlação e transformar esses eixos em variáveis a usar na regressão logística. Os resultados conseguidos mostraram que a análise fatorial e o modelo Logit apresentam bons indicadores na identificação de Bancos com problemas financeiros.

Platt e Platt (1991) utilizaram uma amostra de empresas industriais para comparar modelos com variáveis ajustadas e modelos não ajustados e concluíram sobre o seu uso na previsão de dificuldades financeiras dos mesmos. Neste estudo, os autores pretendiam concluir quais os melhores índices a utilizar neste tipo de questões. Na comparação dos modelos, recorreram a três procedimentos: o teste J para testar a força relativa da especificação do modelo não linear, o teste de Chow para averiguar a estabilidade dos coeficientes da regressão e uma avaliação *ex-ante* para aferir sobre previsão com os dados da amostra. Embora preliminares, os dados mostram que indústrias com índices ajustados são mais significativos e estáveis em comparação com índices não ajustados.

Lawrence e co-autores (1992) analisaram os determinantes de risco no incumprimento do crédito hipotecário. Recorreram ao modelo Logit e usaram como variáveis explicativas o histórico de pagamentos, as condições de aprovação inicial do crédito, as características dos mutuários, as condições económicas e as restrições legais em caso de incumprimento. Este estudo apresentou três grandes contribuições para a questão do risco de crédito: foi o segundo estudo de pesquisa na temática do risco de crédito, complementa estudos anteriores mas modelados com novas variáveis económicas e os dados da amostra são muito representativos pois correspondem a uma carteira de crédito de uma grande instituição bancária. Os autores concluíram que o histórico de pagamentos é um bom indicador para estimar o risco de incumprimento a médio longo prazo.

Smith e Lawrence (1995) aplicaram o método dos mínimos quadrados ordinários (OLS) e o Logit em dados de uma carteira de uma instituição financeira para prever quais as reservas de capital necessário para responder a possíveis perdas. Esta análise permite avaliar o risco de crédito e assim adequar o capital necessário face ao montante concedido. As variáveis usadas foram estratificadas em oito grupos diferentes e os coeficientes de regressão mostram que as variáveis analisadas são bons preditores de incumprimento de crédito. Os métodos utilizados provam ser consistentes mas os autores alertam para os possíveis problemas do método OLS em regressões cuja variável dependente é binária. Nada garante que a probabilidade de incumprimento se situa entre 0 e 1. No caso do método Logit este problema é ultrapassado devido à

consistência associada a este método. Este estudo permitiu à Instituição estudada diminuir as suas reservas perante a carteira de crédito sobre gestão.

Lehmann (2003) estudou se a informação qualitativa era susceptível de melhorar a qualidade de notação de crédito. Assim foi feita uma comparação entre dois modelos: um com informação qualitativa e outro só com informação quantitativa. A amostra foi recolhida num Banco alemão e era composta por 20000 observações de pequenas e médias empresas. A regressão logística foi usada para prever a probabilidade de insolvência. Este estudo mostra que a inclusão de variáveis como a informação qualitativa e os julgamentos subjetivos, são capazes de transmitir informação valiosa e melhorar os sistemas de notação de crédito.

Atendendo ao papel das pequenas e médias empresas no novo acordo de capital, Altman e Sabato (2007) desenvolveram um modelo de previsão de falência específico para as pequenas e médias empresas e analisaram a sua eficácia com o modelo *Z-Score*. Os dados dizem respeito a 2000 empresas dos Estados Unidos da América e foram recolhidos entre os anos 1994 e 2000. A regressão utilizada foi o Logit e o objetivo do trabalho era observar a capacidade deste modelo estimar as necessidades de reduzir o capital dos Bancos, considerando as novas regras do acordo de capital, e a motivação fundamentou-se na importância da modelação do risco de crédito para pequenas e médias empresas em separado das grandes empresas. Os resultados mostram que são necessários modelos e procedimentos específicos para o segmento das pequenas e médias empresas e esta separação demonstra ainda que há mais benefícios em termos de rentabilidades de negócios para os bancos. Os autores referem ainda a necessidade de criar um sistema instrumental de pontuação e classificação neste segmento.

Batista (2012) também escreve sobre a parametrização do risco de crédito. Podemos ver no seu livro que Freed e Glover (1981) utilizaram a programação linear. Breiman e Friedman (1973), Makowski (1985), Coffman (1986), Safavian e Landgrebe (1991) usaram a técnica não paramétrica “Árvores de Decisão” na classificação de indivíduos. Podemos ainda ver as Redes Neurais Artificiais usadas por Makowski (1985) e Coffman (1986) no *Credit Scoring*.

Os modelos de risco de crédito e a sua parametrização tiveram uma evolução significativa na segunda metade do século XX e com resultados ainda hoje válidos quando aplicados. Ao longo da literatura verifiquei que a técnica da análise

discriminante foi a mais usada e os modelos *Z-Score* e *Zeta* de Altman provam os bons resultados dessa técnica.

Outros autores usaram técnicas diferentes para os mesmos fins. Por exemplo Martin (1977), West (1984), Platt e Platt (1991), Lawrence e co-autores (1992), Smith e Lawrence (1995) usaram a regressão logística em complemento das técnicas análise discriminante e análise fatorial.

Mas existem outras mudanças interessantes na especificidade da informação usada. As técnicas anteriores baseiam-se em variáveis de natureza quantitativas mas Lehmann (2003) e Grunert (2004) estudaram o impacto que as variáveis de natureza qualitativa têm nos processos de decisão e concluíram que o seu uso é importante na tomada de decisão.

Assim, o Método Logit e a Análise Discriminante são duas das técnicas estatísticas mais utilizadas no tratamento de dados e no processo de decisão de crédito na segunda metade do século XX (Quadro 1).



#### 4- Descrição das técnicas estatísticas mais utilizadas na literatura

Nesta secção pretendo descrever a importância da análise discriminante e do modelo Logit no uso do tratamento de dados na análise multivariada.

Apesar de não poder ser utilizada na amostra em estudo devido ao número reduzido de empresas insolventes, é de especial interesse verificar em que consiste a análise discriminante e o que permite concluir. Por outro lado o modelo Logit é usado para calcular os coeficientes das variáveis usadas no estudo. Estas serão obtidas pela análise fatorial e depois utilizadas como variáveis independentes no estudo das variáveis dependentes “Situação”, “Incidentes” e “Endividamento”.

##### 4.1- Análise Discriminante

Consultando Marôco (2011), a análise discriminante é uma técnica estatística multivariada que tem por objetivo: i) a identificação das variáveis que melhor distinguem dois ou mais grupos de indivíduos estruturalmente diferentes e mutuamente exclusivos; ii) a utilização dessas variáveis para criar um “índice” ou “função discriminante” que represente de forma parcimoniosa as diferenças entre grupos; iii) a utilização dessa função discriminante para classificar *à priori* novos indivíduos nos grupos.

A seleção das variáveis que melhor discriminam os grupos é feita com recurso à estatística  $\Lambda$  de Wilks (Stevens, 1986),

$$\Lambda_p = \frac{SQE_p}{SQT_p}, \text{ para um } p \geq m$$

em que SQE é a soma dos quadrados dos erros e a SQT a soma dos quadrados totais.

$SQE = \sum_{i=1}^k (n_i - 1) * S_i'^2$ , onde  $S_i'$  é o estimador da variância da amostra  $i$ , e

$SQT = (N-1) * S'^2$ , onde  $S'^2$  é o estimador da variância total.

Considera-se  $p$  o número de variáveis totais iniciais e  $m$  o número de variáveis discriminantes.

A distribuição exata de  $\Lambda$  é dada pela função de distribuição de F-Snedecor, que no caso de apenas dois grupos é,

$$F = \left( \frac{1 - \Lambda}{\Lambda} \right) \left( \frac{N - p - 1}{p} \right)$$

As hipóteses enunciadas para a decisão dessa variável como discriminante são: H0, todas as variáveis têm médias iguais; H1, pelo menos uma média é diferente. A rejeição da H0 leva a concluir que essa variável discrimina bem os grupos.

Dado, p-variáveis e g-grupos em que  $m = \min(g-1, p)$  pode calcular-se a função discriminante ( $D_i$ ) que é dado por

$$D_i = W_{i1}X_1 + W_{i2}X_2 + \dots + W_{ip}X_p, \quad i=1, \dots, m$$

Trata-se de uma função discriminante como combinação linear das p-variáveis e os  $W_{i1}, W_{i2}, \dots, W_{ip}$  representam os pesos das variáveis na função.

Um dos requisitos na análise discriminante é que a variabilidade dos *scores* da função discriminante seja máximo entre os grupos e mínima dentro dos grupos, tal que,

$$\lambda_1 = \left( \frac{SQF(D_i)}{SQE(D_i)} \right)$$

A função discriminante obtida desta forma é conhecida pela Função Discriminante Linear de Fisher.

SQF corresponde à soma dos quadrados do fator e resulta de,

$$SQF = \sum_{i=1}^k n_i (\bar{y}_i - \bar{y})^2$$

Um dos princípios associados ao cálculo das funções discriminantes é que os *scores* das funções discriminantes não estejam correlacionadas com as anteriores e tal forma que se verifique que a  $Cov(D_i, D_j) = 0$

Segundo Reis (2001: Pág. 201), *esta técnica de análise multivariada emprega-se para descobrir as características que distinguem os membros de um grupo dos de outros, de modo que, conhecidas as características de um novo indivíduo, se possa prever a que grupo pertence.*

Assim, a classificação de um novo indivíduo é feita com base nas combinações lineares de variáveis independentes que identificam os grupos à partida. Como conclui Reis (2001: Pág. 205) (...) *a análise discriminante é um método estatístico para classificar indivíduos ou objetos de modo exaustivo em grupos mutuamente exclusivos, com base num conjunto de variáveis independentes. Para isso, são determinadas combinações lineares dessas variáveis que discriminam entre grupos definidos a priori, de tal modo que seja minimizada a probabilidade de erro de incorreta classificação a posteriori.*

## 4.2- Modelo Logit

O Modelo Logit é um método de estimação usado quando a variável endógena é, por exemplo, binária podendo tomar o valor 1 ou 0.

Seja a equação,  $Y_i = \mathbf{X}_i\beta + \mu_i$ , em que no presente estudo,

$$Y_i = \begin{cases} 0, & \text{empresas em atividade} \\ 1, & \text{empresas em insolvência} \end{cases}$$

Pretende-se com esta regressão obter a probabilidade de um acontecimento,

$$\text{Prob}(Y_i=1) = \Lambda(\mathbf{X}_i\beta) = \frac{\exp(\mathbf{X}_i\beta)}{1 + \exp(\mathbf{X}_i\beta)}$$

A estimação dos coeficientes no modelo Logit é executada pelo método da máxima verosimilhança cuja função de verosimilhança é dada por,

$$\mathcal{L} = \prod_0 \text{Prob}(Y_i=0) * \prod_1 \text{Prob}(Y_i=1),$$

Sendo  $\prod_0$  o produto iterado das  $n_0$  observações para o qual  $Y_i=0$  e  $\prod_1$  o produto iterado nas  $n_1$  observações para qual  $Y_i=1$ .

Como se sabe  $Y$  assume apenas dois valores, 0 e 1, então por simplificação,

$$\mathcal{L} = \prod_{i=1}^n (1 - F(\mathbf{X}_i\beta))^{1-Y_i} (F(\mathbf{X}_i\beta))^{Y_i}.$$

O modelo Logit tem também a seguinte função de distribuição logística,

$$\text{Prob}(Y_i=1) = 1/(1 + e^{-\mathbf{X}_i\beta}).$$

O máximo da função logarítmica de verossimilhança é dado pela expressão,

$$\ln \mathcal{L} = \sum_{i=1}^n \{ (1 - Y_i) \ln(1 - F(\mathbf{X}_i\beta)) + Y_i \ln(F(\mathbf{X}_i\beta)) \}$$

E para se aferir os efeitos marginais tem-se,

$$\frac{\partial E(Y_i)}{\partial E(X_{ji})} = f(\mathbf{X}_i\beta) \beta_j$$

O efeito marginal depende do coeficiente associado a  $X_j$  e de um fator de proporcionalidade idêntico,  $f(\mathbf{X}_i\beta)$ , associado a qualquer um dos outros regressores.

Este método de estimação permite contornar três problemas base:

- A não normalidade dos erros, pois estamos perante uma distribuição logística e a variável endógena é binária;
- A existência de heteroscedasticidade pois  $\text{Var}(\mu_i) = E(\mu_i^2) = P_i(1 - P_i)$ ;
- A obrigatoriedade da probabilidade estimada de um acontecimento se situar entre  $[0; 1]$ .

## 5- Descrição da amostra

Este trabalho recai sobre as empresas industriais portuguesas com elementos contabilísticos disponíveis a 31-12-2012 na Base de Dados SABI – Sistema de Análise de Balanços de Empresas Ibéricas.

Atendendo ao elevado número de microempresas e ao reduzido volume de faturação apenas considereei na minha amostra:

- Empresas com o código de atividade económica industrial;
- Empresas com um volume de faturação  $\leq 50$  milhões de euros; e
- Empresas com um número de funcionários compreendido entre 5 e 250.

Esta seleção permite extrair uma amostra total de 18090 empresas.

Neste total tenho presente:

- 440 Empresas PME Excelência 2012 correspondendo a 2,43% da amostra;
- 287 Empresas com pedido de insolvência correspondendo a 1,58% da amostra;
- 17363 Correspondendo à restante amostra.

Pela regra da equivalência, com arredondamento à unidade mais próxima, e para uma amostra de 500 empresas selecionei aleatoriamente:

- 10 Empresas excelência;
- 10 Empresas Insolventes;
- 480 Empresas Industriais.

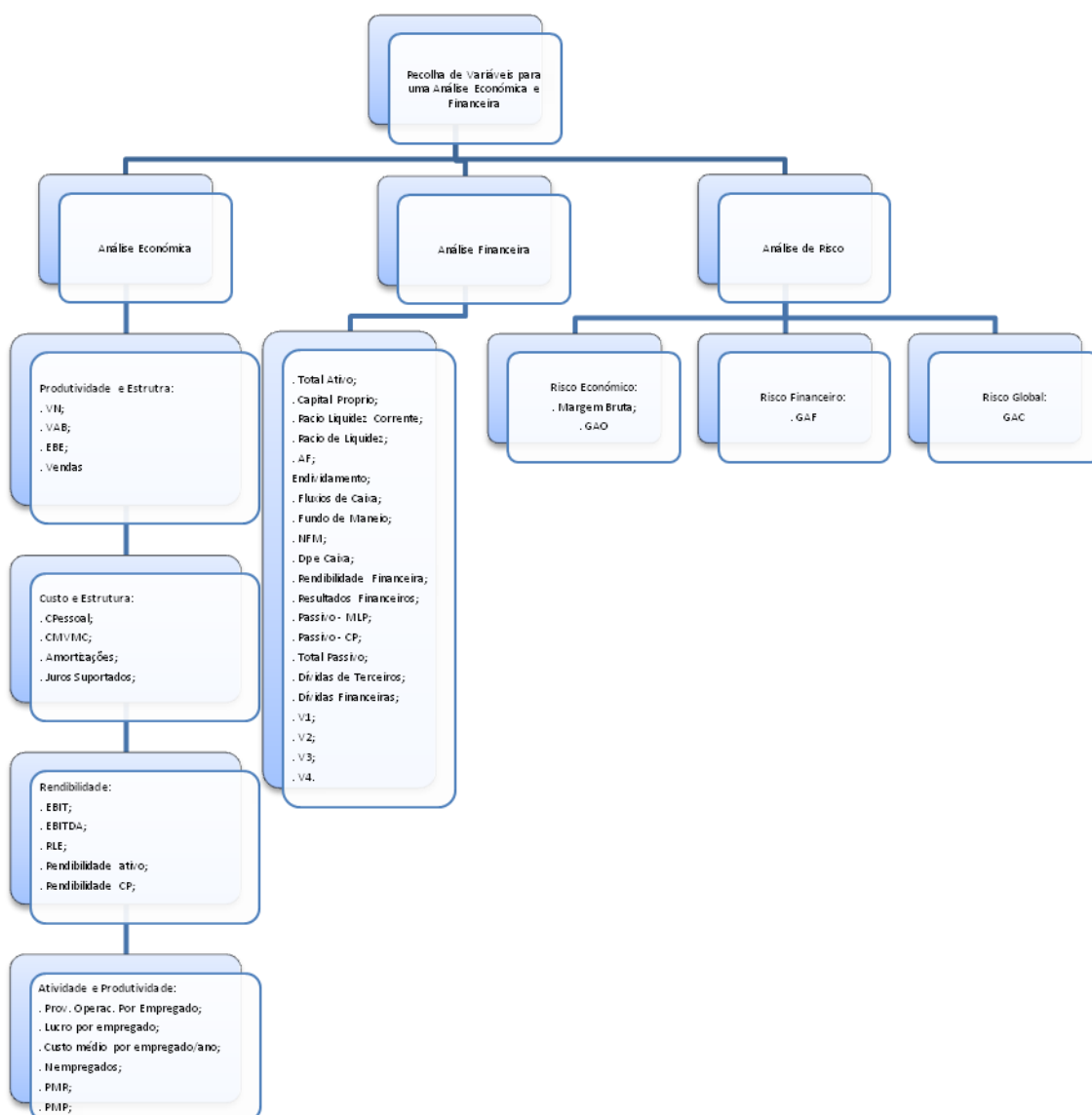
Esta seleção de empresas respeita a regra de proporcionalidade e permite obter uma amostra representativa.

## 6- Definição das variáveis

Segundo Moreira (2001: Pág. 168), *qualquer que seja o objetivo particular do analista, o fim último da análise continuará a ser, sempre, procurar antecipar a evolução futura da empresa.*

A figura 1 esquematiza um conjunto de variáveis selecionadas de acordo com uma análise económica, análise financeira e análise de risco. Todas as variáveis foram retiradas da base de dados SABI e serão a base da análise estatística deste trabalho.

Figura 1: Variáveis selecionadas.



Na análise económica pretende-se mostrar a capacidade da empresa gerar proveitos e suportar os custos. Assim, recolhi variáveis de acordo com quatro especificidades: Produtividade e Estrutura, Custos e Estrutura, Rendibilidade e Atividade e Produtividade.

A Produtividade e Estrutura permitem auferir informação respeitante à quantidade de bens e serviços produzidos e colocados à disposição do consumidor final. Por um lado faz-se a medição via volume de negócios e evolução das vendas através da variação deste indicador e por outro lado calcula-se a produtividade líquida de custos. Isto porque o  $VAB = \text{Produção} - CMVMC - FSE + PSU - Imp$  e o  $EBE = VAB - CP_{\text{Pessoal}}$  permitem retirar os custos materiais e humanos associados ao processo de produção.

Os custos e estrutura têm por finalidade quantificar o peso que a mão-de-obra tem na empresa bem como os custos das matérias-primas e matérias consumidas no processo de fabrico. São ainda recolhidos dados sobre as amortizações inerentes ao ciclo produtivo bem como os juros suportados pelos financiamentos obtidos. Estas rubricas são interessantes porque permitem quantificar os custos e averiguar se estes mesmos custos influenciam os resultados a médio longo prazo.

Na Rendibilidade recolhem-se os resultados líquidos da empresa, os resultados correntes (EBIT) e a capacidade da empresa libertar meios líquidos (EBITDA). Com o resultado líquido consegue-se calcular a rendibilidade do ativo ( $RLE/Ativo$ ) e a rendibilidade do capital próprio ( $RLE/Capital\ Próprio$ ). Este grupo de variáveis é porventura o mais consultado quanto se pretende verificar no imediato e de forma sucinta os resultados das empresas e a sua capacidade de gerar meios líquidos.

Na Atividade e Estrutura calcula-se o peso da mão-de-obra nos resultados finais obtidos nas empresas. Por um lado calcula-se os proveitos operacionais e por outros os custos. É possível assim medir o contributo de cada empregado nos resultados finais. De outra forma consideram-se também o prazo médio de recebimento e o prazo médio de pagamento, ambos quantificados em número de dias. Este conjunto de variáveis permite concluir sobre a eficiência da empresa gerir os recursos humanos e parametrizar o tempo médio de recebimento e pagamento.

Moreira (2001: Pág. 175) refere que a análise financeira (...) *gira à volta do estudo do equilíbrio financeiro, entendendo-se por tal a situação em que a empresa tem capacidade para em cada momento solver os seus compromissos financeiros*. Também

Fernandes e co-autores (2013: Pág. 29) referem que o equilíbrio financeiro recai sobre a capacidade da empresa pagar atempadamente as suas dívidas e ser capaz de solver os seus compromissos a médio e longo prazo. Sinalizam também que a capacidade de solvência dos compromissos está dependente do grau da cobertura dos ativos por parte dos capitais próprios e também da capacidade da empresa gerar resultados positivos.

Para a parametrização desta rubrica recolhi dados do ativo, capital próprio e passivo. Procedi também ao levantamento de informação sobre a autonomia financeira, sobre os rácios de liquidez e fluxos de tesouraria, níveis de endividamento e de terceiros bem como sobre a rentabilidade financeira. Esta panóplia de variáveis irá fornecer informação crucial para uma abordagem sobre a capacidade de solvência da empresa e a sua postura sobre os compromissos assumidos a médio e longo prazo. São também considerados quatros variáveis que resultam de rácios financeiros. A variável V1 mede o peso que o passivo total tem sobre o volume de negócios. A V2 mede o peso do passivo total em relação aos meios libertos líquidos. A V3 representa a autonomia financeira e a variável V4 mede a capacidade de reembolso ((passivo total – ativo)/(resultados líquidos + amortizações)).

Segundo Fernandes e co-autores (2013: Pág. 238) o risco económico deve ser visto como, *a probabilidade do resultado operacional (EBIT) ser inadequado aos objetivos da empresa, ou seja, que os rendimentos ou ganhos operacionais não sejam suficientes para cobrir os gastos ou perdas operacionais. (...) assim pode entender-se o risco económico como a incerteza inerente à realização e continuidade de resultados operacionais futuros*. Nesta rubrica recolhi dados sobre as variáveis Margem Bruta e GAO. A Margem Bruta=Vendas – Gastos variáveis. E o GAO= $[(\Delta RO/RO)/(\Delta V/V)]$ . O GAO é um indicador de risco que procura medir a volatilidade dos resultados operacionais face a variações ocorridas ao nível das vendas.

Segundo Moreira (2001: Pág. 114) o risco financeiro, *respeita à incerteza inerente à projeção da rentabilidade dos capitais próprios, devido à utilização de financiamento alheio, que induz a existência de encargos financeiros fixos. Quanto maior for o peso dos capitais alheios na estrutura de capitais, maior tenderá a ser este tipo de risco*. O indicador utilizado é o GAF =  $[(\Delta RL/RL)/(\Delta RO/RO)]=(RO/RAI)$ . Este indicador mede a sensibilidade do resultado líquido da empresa a variações nos resultados operacionais.

Segundo Fernandes e co-autores (2013: Pág. 246) *a conjugação da análise do risco económico ou operacional com o estudo do risco financeiro permite fazer uma avaliação do risco global da empresa. Esta avaliação terá mais interesse e utilidade quando comparados os valores das suas componentes ao longo do tempo e feita a sua confronto com os dados setoriais.* O  $GAC = GAO * GAF = (MB/RAI)$ . Desta forma o GAC mede a sensibilidade do resultado líquido corrente às variações do volume de vendas.

## 6.1- Estatísticas descritivas das variáveis

Atendendo ao elevado número de variáveis em estudo, neste subcapítulo analiso as variáveis: “Endividamento”, “Autonomia Financeira”, “Prazo Médio de Recebimento”, “Prazo Médio de Pagamento”, “Fundo de Maneio”, “Rendibilidade Ativo”, “Rendibilidade do Capital Próprio” e “Rendibilidade Financeira”. Este conjunto de variáveis permite aferir qual a exposição das empresas ao financiamento, a sua gestão de pagamentos, do fundo de maneio e a obtenção de rendibilidades.

Em termos médios (quadro 2), as empresas da amostra apresentam um nível de endividamento de 81,20% e uma autonomia financeira de 17,74%. Estas duas variáveis mostram uma elevada exposição ao financiamento e uma debilidade na estrutura de capitais próprios que se refletem na autonomia financeira.

Quadro 2: Estatísticas descritivas das variáveis da amostra.

		Endividamento t	AF t	PMR t	PMP t	Fundo de maneio t	Rendibilidade Ativo t	Rendibilidade CP t	Rendibilidade Financeira t
N	Válido	500	500	500	500	500	500	500	500
	Ausente	0	0	0	0	0	0	0	0
Média		<b>,811970</b>	<b>,17739</b>	<b>131,14</b>	<b>85,81</b>	<b>463,5390</b>	-,052279	-1,597743	-,522325
Mediana		,853050	,09401	105,84	57,04	79,1888	<b>,004244</b>	<b>,046195</b>	<b>,042700</b>
Assimetria		5,708	-,346	1,846	2,871	4,491	-3,191	-8,183	-1,901
Erro de Assimetria		,109	,109	,109	,109	,109	,109	,109	,109
Curtose		72,494	8,971	4,082	11,13	25,405	26,300	87,340	7,350
Erro de Curtose		,218	,218	,218	,218	,218	,218	,218	,218
Mínimo		,0439	-1,7645	-11	1	-214,89	-2,0422	-100,0173	-9,6260
Máximo		5,8558	,9533	665	760	9134,20	,8638	9,2980	7,3711



Os prazos de pagamentos e de recebimentos situam-se, em termos médios, nos 86 e 131 dias, respetivamente. O saldo médio do fundo de maneio é de 464 euros.

Analisando as rendibilidades, em termos médios, estas tendem a ser negativas. Considerando apenas os valores centrais da rendibilidade do ativo, do capital próprio e financeira, os valores obtidos são 0,4%, 4,61% e 4,27% respetivamente.

Dando enfoque ao coeficiente de assimetria, os resultados apresentados tendem a evidenciar que a distribuição dos dados não é normal.

## 7- Análise em componentes principais (ACP): a importância no tratamento de dados

Consultando Marôco (2011), a análise de componentes principais é uma técnica de análise exploratória multivariada que transforma um conjunto de variáveis correlacionadas num conjunto menor de variáveis independentes que são combinações lineares das variáveis originais e definem-se como componentes principais. A análise de componentes principais é geralmente encarada como um método de redução da complexidade dos dados.

Algebricamente, seja  $p$  um conjunto de variáveis iniciais correlacionadas entre si e pretende-se obter  $p$  combinações lineares independentes, então segundo Johnsen & Wichern (1992) temos,

$$\xi_1 = \gamma_{11}X_1 + \gamma_{12}X_2 + \dots + \gamma_{1p}X_p$$

(...)

$$\xi_p = \gamma_{p1}X_1 + \gamma_{p2}X_2 + \dots + \gamma_{pp}X_p, \text{ sendo}$$

$X_1, X_2, \dots, X_p$  o conjunto de variáveis e as componentes principais dadas por  $\xi_1, \dots, \xi_p$ . Considera-se  $\gamma_{ij}$  o peso que cada variável tem nessa componente principal.

A 1ª componente principal é responsável pela maior explicação da variância das variáveis iniciais. A 2ª componente principal é responsável pela maior capacidade de explicação da variância não explicada na 1ª componente principal e garante a independência da componente anterior. Cada componente adicional garante independência da anterior e consequentemente o número de componentes obtido garante constante a variância total de modo a que,

$$\gamma_{11}^2 + \gamma_{12}^2 + \dots + \gamma_{1p}^2 = 1$$

Assim, este método transforma variáveis correlacionadas em novas variáveis não correlacionadas, diminui os dados da amostra e conduz a uma simplificação da informação para uma leitura mais simples dos dados.

## 7.1- Aplicação da ACP aos dados da amostra: distribuição de variáveis e indivíduos

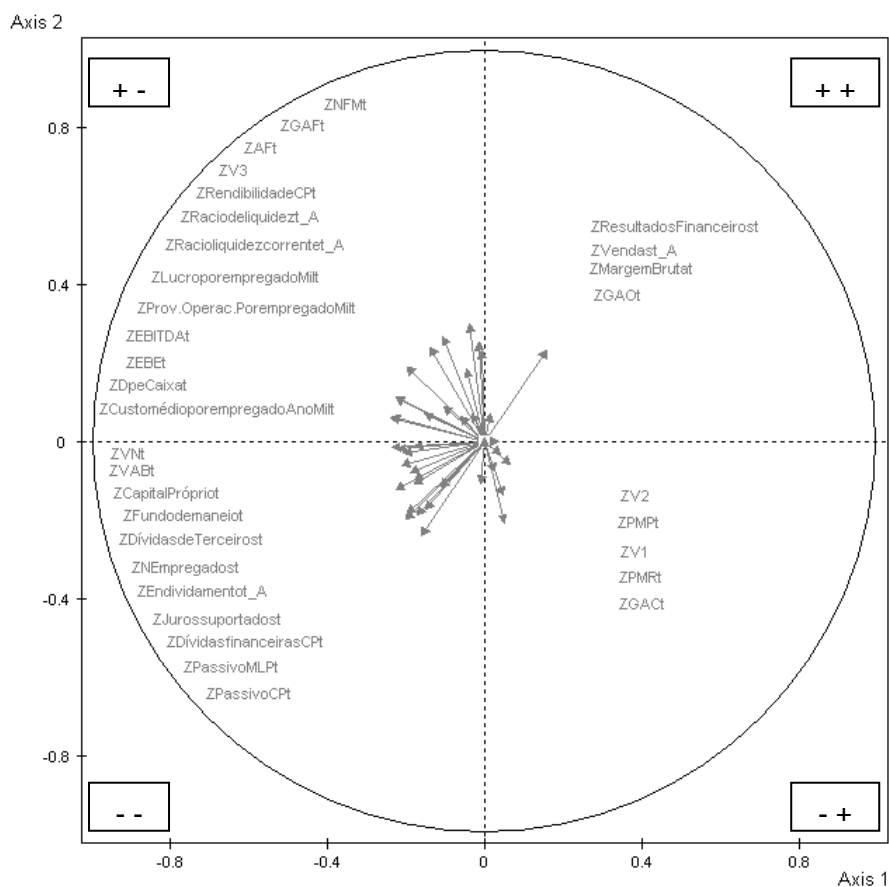
A ACP foi aplicada aos dados da nossa amostra para perceber como se organizam as variáveis em grupos e também como se distribuem e se agrupam os indivíduos através das suas características.

Foram tidos em consideração dois pressupostos iniciais:

- 1) Esta técnica de agrupamento de informação foi aplicada às variáveis de natureza quantitativa;
- 2) As variáveis foram estandardizadas de modo a que a variância entre as variáveis seja exatamente igual em toda a amostra;

Os resultados obtidos encontram-se representados nas figuras 2 e 3 e foram efetuadas com recurso ao software SPAD.

Figura 2: Distribuição das Variáveis.



A figura 2 representa graficamente a distribuição das variáveis de natureza quantitativa usadas no estudo nos quatro quadrantes (++), (+-), (--), e (-+).

A distribuição das variáveis ao longo dos quatro quadrantes mostra que não se parte de um grupo de variáveis previamente definido e dividido na análise económica, financeira e de risco como mostra a figura 1.

O quadrante (++) mede os resultados financeiros, a margem bruta e a volatilidade dos resultados operacionais a variações das vendas. Analisando as quatro variáveis que compõem este quadrante conclui-se que, tal como se pretende com a ACP, existe um fator comum entre as variáveis que neste caso são as vendas. Fazendo uma interpretação a cada uma das variáveis verifica-se que os resultados financeiros dependem da atividade da empresa que está relacionada com o resultado do lucro nas vendas. No que respeita à margem bruta esta depende das vendas e dos gastos variáveis da empresa. E por fim o GAO depende da volatilidade dos resultados operacionais a variações ocorridas nas vendas.

No quadrante (+-) agrupam as variáveis associadas aos resultados obtidos pela empresa, os rácios de rendibilidade e a sensibilidade dos resultados líquidos perante os resultados operacionais bem como a diluição do lucro e dos custos pela mão-de-obra. Assim concluo que o fator comum transversal a estas variáveis é o resultado líquido obtido pela empresa

No quadrante (--) estão agrupadas as variáveis associadas ao passivo, às dívidas financeiras, custo de financiamento, estrutura do capital próprio e volume de negócios. Esta informação leva a concluir que um possível fator subjacente neste quadrante é o financiamento da existente na empresa.

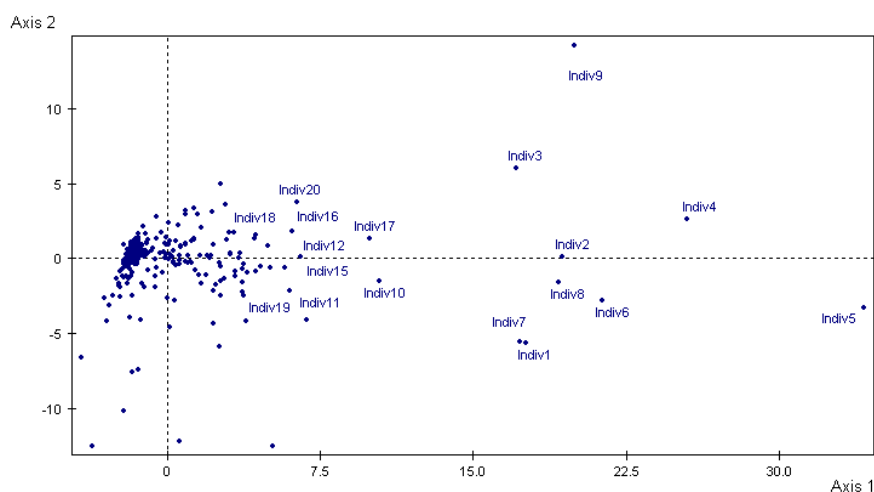
Por fim no quadrante (-+) estão presentes as variáveis associadas aos prazos de pagamentos, ao “peso” do passivo sobre o volume de negócios e o passivo sobre os resultados correntes. Perante a natureza das variáveis que compõem este quadrante um possível fator comum está associado aos prazos de pagamento e de recebimento. Este fator mede o número de dias que a empresa demora a pagar e a receber e que por sua vez pode influenciar o passivo e a sua parametrização ao nível dos resultados correntes.

Esta seletividade de variáveis permite, numa primeira abordagem, verificar o modo como as variáveis se vão posicionar aquando a sua extração na análise fatorial. Os resultados obtidos pela análise em componentes principais permitem concluir que é

possível reduzir a informação inicial em componentes que mantêm a informação original mas torna a sua leitura mais simples. O desafio na análise fatorial será extrair o menor número de fatores mantendo o máximo de informação inicial e utilizá-los, posteriormente, na regressão de modelos econométricos, Logit e OLS, como variáveis explicativas.

A figura 3 apresenta a distribuição das empresas ao longo do plano 1-2. O grupo de indivíduos mais desfaçados à direita diz respeito às empresas industriais com estatuto PME Excelência em 2012. Ao nível central encontram-se as empresas classificadas como insolventes e à esquerda, e em maior número, posicionam-se as restantes empresas industriais. Esta distribuição dos indivíduos evidencia não só a heterogeneidade da informação mas também a capacidade de extração que a análise fatorial terá de ter para reduzir a informação mantendo o essencial da informação original.

Figura 3: Distribuição dos Indivíduos.



Feito o levantamento dos dados anteriormente apresentados é importante referir que, segundo Reis (2001: Pág. 261), *Na análise das componentes principais, a representação matemática das combinações lineares não pressupõe a imposição de qualquer modelo causal mas também não permite detetar qualquer relação de causa-efeito entre as variáveis iniciais mesmo que existam.*

Assim, a secção seguinte irá recair sobre a análise fatorial. Tem-se como princípio a aferição de uma parte da informação associada a cada variável e uma outra parte comum a todas as outras variáveis. A análise fatorial tem assim como base a

decomposição de cada variável em duas partes: uma parte comum e uma parte única. A parte comum tem uma variação partilhada com as outras variáveis e a parte única é específica à sua própria variação.

Como se conclui em Reis (2001: Pág. 262) (...) *uma distinção importante entre os dois métodos provém do montante de variação analisada: enquanto que na análise de componentes principais se considera a variação total presente no conjunto das variáveis originais, na análise fatorial só é retida a variação comum, partilhada por todas as variáveis.*

## 8- Análise fatorial

### 8.1- Descrição da técnica

Como se pode verificar em Maroco (2011) a análise fatorial é uma técnica de análise de dados que tem por finalidade descobrir e analisar a estrutura de um conjunto de variáveis inter-relacionadas de forma a construir uma escala de medida por fatores comuns que de alguma forma controlam as variáveis iniciais.

Assim, esta análise usa as correlações entre as variáveis iniciais para estimar os fatores comuns e as relações estruturais que ligam os fatores às variáveis. É uma técnica de exploração de dados multivariada e o principal objetivo é atribuir um score a fatores que não são diretamente observados. O score atribuído é uma representação parcial da informação presente nas diferentes variáveis sendo capaz de resumir a informação presente em muitas variáveis num menor número de fatores.

Algebricamente, seja o modelo fatorial dado por (Sharma, 1986; Jonhson, 1998):

$Z = \Lambda f + \eta$  em que:

$Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{pmatrix}$  é o vetor das  $p$  variáveis estandardizadas

$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{pmatrix}$  é o vetor dos fatores comuns

$\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_p \end{pmatrix}$  é o vetor dos fatores específicos, e

$\Lambda = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1m} \\ \vdots & \ddots & \vdots \\ \lambda_{p1} & \cdots & \lambda_{pm} \end{bmatrix}$  a matriz dos pesos fatoriais

Sendo o modelo ortogonal, então:

1.  $f$  tem média 0 e variância I;
2.  $\eta$  possui média 0 e variância  $\Psi$ ;
3.  $f$  e  $\eta$  são independentes.

Considera-se I a matriz identidade,

$I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$  e

$\Psi$  uma matriz diagonal,

$$\Psi = \begin{bmatrix} \psi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \psi_p \end{bmatrix}$$

Por fim,

A matriz das correlações,  $\Pi$ , é modelada por:

$$\begin{aligned} \Pi &= E(ZZ') = E[(\Lambda f + \eta)(\Lambda f + \eta)'] = E[(\Lambda f + \eta)(\Lambda' f' + \eta')] = E(\Lambda f f' \Lambda') + E(\eta \eta') \\ &= \Lambda \Lambda' E(f f') + E(\eta \eta') \\ &= \Lambda \Lambda' V(f) + V(\eta) \\ &= \Lambda \Lambda' + \Psi \end{aligned}$$

A análise fatorial parte, portanto, do vetor das variáveis estandardizadas para estimar os pesos dos fatores comuns e específicos. Após estes cálculos determina-se a matriz das correlações. É desta forma interessante que a matriz das correlações ( $\Pi$ ) seja diferente da matriz identidade ( $I$ ) de forma às variáveis estarem muito inter-relacionadas e assim os fatores específicos explicarem a menor proporção da variância e os fatores comuns terem uma maior capacidade explicativa.

## 8.2- Aplicação na amostra

A extração dos fatores será feita pelo método das componentes principais, onde o que pretendo é uma redução da complexidade da amostra e manter uma ou mais combinações linearmente independentes das variáveis correlacionadas com os eixos ou fatores selecionados.

Na decisão do número de eixos a reter recorro a 3 critérios:

1) Considero a representação gráfica da variância explicada por cada componente (Scree plot proposto por Cattell (1966)). Neste critério analisa-se a variância explicada por cada componente principal e considera-se um ponto de corte numa quebra significativa de variância explicada entre duas componentes.

2) Apenas considero as componentes em que o valor próprio é superior a 1. Segundo Reis (2001: Pág. 273) *excluir as componentes cujos valores próprios são inferiores à média, isto é, menores que 1 se a análise for feita a partir de uma matriz de correlações (Critério de Kaiser (1958)).*



3) Por fim, considero o critério de Pearson que consiste em reter os eixos em que a soma da inércia explicada é equivalente a 80%.

Este último critério é interessante porque podem existir vários eixos em que se verificam um valor próprio superior à unidade mas em termos de capacidade explicativa não ser relevantes devido à baixa inercia explicada.

Para o cálculo da matriz ortogonal utilizo o método Varimax. Segundo Reis (2001) é o método mais popular e também o melhor na rotação das componentes principais. É um método proposto por Kaiser (1958) e tem como finalidade maximizar a variação entre os pesos de cada componente principal. Este método ortogonal pretende que, para cada componente principal, existam apenas alguns pesos significativos e os outros sejam próximos de zero.

A rotação é feita através de um processo iterativo de maximização de uma função quadrática (Q) do tipo:

$$Q = \sum_{j=1}^p \{ \sum_{i=1}^p b_{ij}^4 - (c/p) (\sum_{i=1}^p b_{ij}^2)^2 \}$$

Em que,

b = vetor;

q = número de componentes principais;

c = é a constante;

p = é o número das variáveis.

Exemplificando,

Seja G a nossa matriz ortogonal e considere-se apenas duas componentes principais (q = 2), então G é dada por:

$$G = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

O que representa uma rotação dos eixos equivalentes a um ângulo  $\theta$  no sentido dos ponteiros do relógio.

A proporção de variância explicada por cada uma destas componentes mantém-se constante e apenas se distribui de um modo diferente para que sejam maximizadas as diferenças entre as contribuições das variáveis. Aumentam as que mais contribuem para a formação da componente e diminuem os pesos das que menos contribuem.

Segundo Marôco (2011) esta rotação de fatores tem por objetivo produzir uma solução interpretável. Neste método uma variável inicial apenas está fortemente correlacionada com um fator e pouco correlacionada com os restantes.

A estimação dos valores dos fatores é feita pelo método dos mínimos quadrados ponderados, também conhecido pelo método de Bartlett.

Consultando Marôco (2011), neste método, depois de definido o modelo  $Z = \Lambda f + \eta$ , os *scores*  $f_i$  para o indivíduo  $i$  podem ser determinados por minimização da função,

$$(Z_i - \Lambda f) \Psi^{-1} (Z_i - \Lambda f),$$

Cuja solução ocorre para,

$$f_i = (\Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} Z_i \text{ em que } i=1, \dots, n$$

Na avaliação da qualidade do modelo recorri ao *Goodness of Fit Index* (GFI) normal (Mulak *et al.*, 1989) e completei o resultado com um ajustamento aos graus de liberdade. O indicador *Root Mean Square Residual* (RMSR) também foi utilizado para reforçar a validade da qualidade do modelo.

O recurso a estes índices de avaliação prende-se com o facto de aferir se a variância total explicada reproduz bem as correlações entre as variáveis originais e se a estrutura fatorial reduzida se ajusta aos dados.

$$\text{O índice (GFI)} = 1 - [(\text{tr}[\mathbf{W}^{-0.5}(\mathbf{S} - \Sigma(\theta))\mathbf{W}^{-0.5}]^2) / (\text{tr}[\mathbf{W}^{-0.5}\mathbf{S}\mathbf{W}^{-0.5}]^2)]$$

Dado,  $\mathbf{W} = \mathbf{S}$ , então

$$\text{GFI} = 1 - 0,5 \text{tr}(\mathbf{R} - \mathbf{\Pi})^2$$

Em que,

$\text{tr}$  = é a função traço (corresponde à soma dos elementos da diagonal principal da matriz);

$\mathbf{R}$  = matriz das correlações observadas;

$\mathbf{\Pi}$  = matriz das correlações estimadas.

Critérios de avaliação:

Se  $\text{GFI} > 0,9$  então o modelo ajusta-se razoavelmente bem aos dados;

Se  $\text{GFI} > 0,95$  então o modelo apresenta um ajustamento muito bom.

Por forma a corrigir possíveis sobrestimações do valor de ajustamento, procedi ao cálculo do GFI ajustado (Schumacker & Lomax, 1996):

$$\text{AGFI} = 1 - (K/df)(1 - \text{GFI}),$$

cujos critérios de avaliação são idênticos ao anterior.

No que respeita ao índice de avaliação RMSR este é calculado a partir do quadrado médio entre as diferenças das variâncias-covariâncias estimadas e as variâncias-covariâncias observadas (Shumacker e Lomax, 1996):

$$\text{RMSR} = [(\sum_{i=1}^p * \sum_{j=1}^p (\sigma_{ij} - s_{ij})^2) / (p(p+1)/2)]$$

Considerando a estandardização das variâncias, tem-se a função (Schumacker e Lomax, 1996; Meehl e Wallee, 2002):

$$\text{RMSR}^* = [(\sum_{i=1}^{p-1} * \sum_{j=i+1}^p (\rho_{ij} - r_{ij})^2) / (p(p-1)/2)]$$

O critério de avaliação é:

$\text{RMSR}^* > 0,1$ : o ajustamento é inaceitável;

$\text{RMSR}^* < 0,1$ : o ajustamento é bom;

$\text{RMSR}^* < 0,05$ : o ajustamento é muito bom.

### 8.3- Pressupostos iniciais

A análise fatorial foi aplicada em 44 variáveis que compõem a amostra tendo como referência o ano de 2012 (ano t). Os dados foram estandardizados, foi considerada a matriz das correlações e os fatores foram extraídos pelo método das componentes principais. A rotação dos fatores foi feita pelo método Varimax com normalização de Kaiser, os *Scores* de cada sujeito em cada uma das componentes principais foram obtidos pelo método de Bartlett e a qualidade do modelo foi avaliada pelos índices GFI, AGFI e RMSR.

Todas estas análises foram efetuadas com o software SPSS Statistics (V.21, IBM SPSS, Chicago, IL) e as componentes retidas foram, posteriormente, utilizadas como variáveis na regressão logística, linear e feita a sua inferência estatística.

### 8.4- Resultados obtidos com a análise fatorial

A tabela 1 mostra que a estrutura relacional das variáveis é explicada por 12 fatores latentes. Considera-se como critérios de seleção dos fatores os que apresentam um valor próprio superior a 1 e a capacidade de inércia explicada até 80%. Assim, a tabela 1 permite verificar que foram extraídos 12 componentes principais, ou fatores, que são responsáveis por 80,17% da variância total explicada. À medida que os fatores foram

extraídos, a variância individual explicada foi diminuindo. Por exemplo, passou de 36,10% do primeiro fator para 2,26% para o fator 12. Outra observação prende-se ao valor próprio de cada componente que sendo sempre superior a 1 passa de 16,61 do fator número 1 para 1,04 no fator número 12.

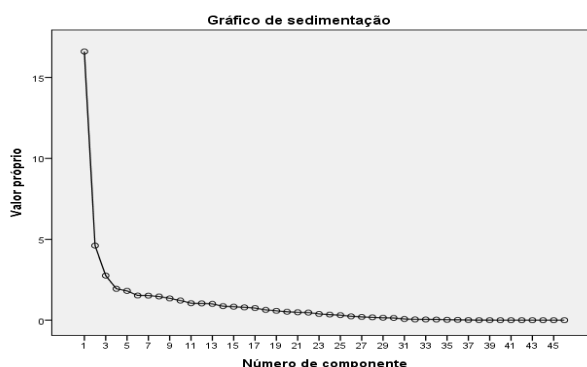
Tabela 1: Variância total explicada.

Variância total explicada									
Componente	Valores próprios iniciais			Somadas de extração de carregamentos ao quadrado			Somadas rotativas de carregamentos ao quadrado		
	Total	Variância (%)	Percentagem cumulativa	Total	Variância (%)	Percentagem cumulativa	Total	Variância (%)	Percentagem cumulativa
1	16,607	36,103	36,103	16,607	36,103	36,103	11,719	25,476	25,476
2	4,611	10,023	46,126	4,611	10,023	46,126	8,004	17,399	42,875
3	2,753	5,986	52,111	2,753	5,986	52,111	2,722	5,918	48,794
4	1,941	4,22	56,331	1,941	4,22	56,331	2,011	4,372	53,166
5	1,813	3,941	60,272	1,813	3,941	60,272	1,951	4,241	57,407
6	1,524	3,312	63,584	1,524	3,312	63,584	1,919	4,172	61,579
7	1,516	3,296	66,88	1,516	3,296	66,88	1,604	3,487	65,066
8	1,464	3,184	70,063	1,464	3,184	70,063	1,564	3,401	68,467
9	1,347	2,928	72,991	1,347	2,928	72,991	1,561	3,393	71,859
10	1,215	2,641	75,632	1,215	2,641	75,632	1,484	3,226	75,085
11	1,049	2,281	77,913	1,049	2,281	77,913	1,179	2,562	77,648
12	1,038	2,256	80,169	1,038	2,256	80,169	1,117	2,428	80,075

Método de extração: análise do componente principal.

Analisando a distribuição dos fatores (Gráfico 4) não há uma quebra significativa de variância entre os fatores ao longo da sua extração. Optar pelo critério de retenção dos fatores cujo valor próprio é superior a 1 e a capacidade de variância explicada de 80% é assim o critério mais indicado.

Gráfico 4- Gráfico de sedimentação.



A tabela 2 mostra pesos fatoriais elevados nas variáveis: Juros Suportados, Total do Passivo, Total do Ativo, Passivo-curto prazo, Custos com o Pessoal, Dívidas Financeiras de curto prazo, Fundo de Maneio, Amortizações, Dívidas de Terceiros, Nº de Empregados, VAB, Capital Próprio, VN, Passivo-mlp e Resultados Financeiros. Todas as variáveis apresentam valores elevados ( $>0,50$ ) na coluna da comunalidade demonstrando que o fator é apropriado para descrever a estrutura correlacional entre as variáveis. Em termos individuais, à exceção do Passivo-Mlp, todas as variáveis estão bem representadas nos eixos com uma contribuição superior a 0,50. Este fator apresenta um valor próprio de 16,607 e é responsável por uma variância explicada de 36,10% (tabela1).

Este fator opõe variáveis de valores a pagar (ex. dívidas financeiras –cp) e a receber (ex. dívidas de terceiros), bem como indicadores de atividade (ex. VN) face à variável Resultados Financeiros.

Tabela 2: “Encargos e Responsabilidades”

Componente Principal 1 - "Encargos e Responsabilidades"				
Valor Próprio	16,607			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore: Juros suportados t	,888		0,789	0,829
Zscore: Total Passivo t	,888		0,789	0,932
Zscore: Total Activo t	,885		0,784	0,964
Zscore: Passivo-CP t	,878		0,771	0,854
Zscore: Cpessoal t	,833		0,693	0,866
Zscore: Dívidas financeiras-CP t	,815		0,665	0,772
Zscore: Fundo de maneio t	,812		0,659	0,876
Zscore: Amortizações t	,791		0,626	0,745
Zscore: Dívidas de Terceiros t	,785		0,616	0,840
Zscore: N Empregados t	,782		0,612	0,771
Zscore: VAB t	,777		0,603	0,960
Zscore: Capital Próprio t	,731		0,534	0,884
Zscore: VN t	,714		0,509	0,923
Zscore: Passivo-MLP t	,554		0,307	0,760
Zscore: Resultados Financeiros t		-,853	0,727	0,773
Componente dos "Encargos e Responsabilidades"				

Todas estas variáveis influenciam os resultados financeiros e, de acordo com a representatividade das variáveis, define-se esta componente como o fator dos “Encargos e Responsabilidades”.

O fator 2 presente na tabela 3 tem um valor próprio de 4,611 e uma capacidade explicativa de 10,0% (ver tabela 1). É constituído pelas variáveis associadas aos resultados do exercício. O método fatorial é representativo na explicação das variáveis por estas apresentarem valores superiores a 0,50, veja-se a coluna das comunalidades. O nível de correlação das variáveis no eixo é elevado, com uma exceção na rubrica “Dp e caixa t”. Em termos individuais as variáveis estão bem representadas no fator pois apresentam-se com uma contribuição superior a 0,50.

Tabela 3: “Resultados Obtidos”

Componente Principal 2 - "Resultados Obtidos"				
Valor Próprio	4,611			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore: RLE t	,960		0,923	0,973
Zscore: EBIT t	,898		0,807	0,970
Zscore: MLL t	,847		0,717	0,956
Zscore: Fluxos de caixa t	,845		0,715	0,957
Zscore: EBITDA t	,804		0,647	0,979
Zscore: EBE t	,797		0,635	0,978
Zscore: Lucro por empregado (Mil) t	,679		0,461	0,819
Zscore: Dp e Caixa t	,572		0,327	0,588
Componente dos "Resultados Obtidos"				

Atendendo à natureza das variáveis que compõem este grupo (resultados obtidos, libertação de meios e saldos de depósitos bancários e caixa) pode concluir-se que o fator que as une é um indicador de liquidez. Este fator é designado como o fator dos “Resultados obtidos”.

O fator 3 opõe a variável Autonomia Financeira ao nível de Endividamento (ver tabela 4). Ambas as variáveis estão bem representadas no eixo, apresentam um elevado nível de correlação no fator e individualmente tem um contributo superior a 50%.

Tabela 4: “Solvabilidade”

Componente Principal 3 – “Solvabilidade”				
Valor Próprio	2,753			
Variáveis	(+)	(-)	C.Individual	Comunalidades
Zscore: AF t	,892		0,796	0,907
Zscore: Endividamento t		-,816	0,666	0,760
Componente da "Solvabilidade"				

A combinação destas duas variáveis é responsável por 5,99% de variância explicada (ver tabela 1) e o seu valor próprio é de 2,753.

A autonomia financeira mede a independência da empresa ao capital alheio e o endividamento mede a dependência da empresa ao capital alheio. Assim o fator entre estas duas variáveis diz respeito à relação da empresa ao capital alheio.

A informação presente no fator nº3 pode ser interpretada como o fator da “Solvabilidade”.

Tabela 5: “Liquidez”

Componente Principal 4 – “Liquidez”				
Valor Próprio	1,941			
Variáveis	(+)	(-)	C.Individual	Comunalidades
Zscore: Racio liquidez corrente t	,974		,974	0,966
Zscore: Racio de liquidez t	,970		,970	0,958
Componente da “Liquidez”				

O fator nº4 (ver tabela 5), apresenta os dois rácios de liquidez. É responsável por 4,22% da variância explicada e o valor próprio da componente é de 1,941 (ver tabela 1). Em termos gerais, a presença das duas variáveis neste fator é significativo porque tanto ao nível de extração da informação como ao nível de representatividade na componente, os valores são elevados.

Estes dois indicadores medem o equilíbrio financeiro dos recursos, que representam o valor imediatamente disponível na empresa, para solver os compromissos de curto prazo, que por definição se define como o passivo corrente. Assim, um fator comum nestas variáveis é o grau de liquidez na cobertura do passivo de curto prazo.

Este fator é assim conhecido como o fator da “Liquidez”.

O fator nº5 (tabela 6) é responsável pelos proveitos tanto ao nível dos recursos humanos como ao nível das matérias-primas utilizadas no processo de fabrico. Esta componente é responsável por 3,94% da variância total explicada e o seu valor próprio é de 1,813. As variáveis são individualmente significativas de acordo com o nível de correlação. A rubrica dos proveitos operacionais tem um contributo individual superior ao CMVMC e ambos contribuem para que este fator se designe o fator dos “Proveitos Operacionais Correntes”.

Um fator comum entre as duas variáveis é os proveitos operacionais como resultados da gestão dos custos variáveis das empresas.

Tabela 6: “Proveitos Operacionais Correntes”

Componente Principal 5 – “Proveitos Operacionais Correntes”				
Valor Próprio	1,813			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore: Prov. Operac. Por empregado (Mil) t	,853		0,728	0,927
Zscore: CMVMC t	,644		0,415	0,925
Componente dos "Proveitos Operacionais Correntes"				

O fator nº6 é conhecido como o fator da “Rendibilidade” (ver tabela 7). A rendibilidade financeira e do ativo está bem representada na componente e todas as variáveis foram extraídas corretamente pelo método fatorial. Esta componente é responsável por uma variância de 3,31% e o seu valor próprio é de 1,524.

Tabela 7: “Rendibilidade”

Componente Principal 6 – “Rendibilidade”				
Valor Próprio	1,524			
Variáveis	(+)	(-)	C.Individual	Comunalidades
Zscore:Rendibilidade Financeira t	,791		0,625	0,65
Zscore: Rendibilidade Ativo t	,771		0,595	0,683
Zscore: Rendibilidade CP t	,496		0,246	0,551
Componente da "Rendibilidade"				

Os rácios de rendibilidade medem a capacidade das empresas gerarem lucros. A rendibilidade do ativo mede o desempenho dos capitais totais investidos pela empresa e a rendibilidade do capital próprio representa o retorno dos investimentos efetuados pela empresa. Por fim, a rendibilidade financeira mede o desempenho dos lucros em função dos resultados financeiros. Desta forma um fator comum neste grupo de variáveis é a capacidade das empresas obterem lucros.

Como se verifica na tabela 8, o fator nº7 transmite informação sobre o prazo médio de recebimento e de pagamento. A extração da informação das variáveis é feita a um nível superior a 50% o que denota que o fator é importante na explicação das



variáveis extraídas. Este fator tem uma variância explicativa de 3,296 (tabela 1) e um valor próprio de 1,516.

Tabela 8: “Prazos de Pagamentos”

Componente Principal 7 – “Prazos de Pagamentos”				
Valor Próprio	1,516			
Variáveis	(+)	(-)	C.Individual	Comunalidades
Zscore: PMR t	,820		,672	0,718
Zscore: PMP t	,779		,607	0,687
Componente dos "Prazos de Pagamentos"				

O PMR e PMP representam o número de dias que decorrem desde a disponibilização do produto final ao comprador e o recebimento do seu preço bem como o número de dias que a empresa leva a pagar o serviço solicitado. Desta forma, estas duas variáveis têm em comum o número de dias que as empresas demoram a converter em meios monetários as suas vendas e a pagarem a suas encomendas.

Este fator é designado pelo fator dos “Prazos de Pagamento”.

O fator nº8 extrai corretamente a informação das duas variáveis, como se vê na tabela 9. Contudo apesar da variável “Custo médio por empregado/ano” ter uma contribuição individual inferior a 50% as duas variáveis são responsáveis por uma variância explicativa de 3,184 e um valor próprio de 1,464.

Tabela 9: “Custos Operacionais Correntes”

Componente Principal 8 – “Custos Operacionais Correntes”				
Valor Próprio	1,464			
Variáveis	(+)	(-)	C.Individual	Comunalidades
Zscore(V1)	,817		0,667	0,766
Zscore: Custo médio por empregado / Ano (Mil) t	,599		0,359	0,646
Componente dos "Custos Operacionais Correntes"				

A rubrica V1 mede o peso que o Passivo Total tem sobre o Volume de Negócios. Esta variável em conjunto com o custo médio por empregado leva a definir este fator como o fator dos “Custos Operacionais Correntes”. Estas duas variáveis representam os custos que as empresas têm diluído pelos recursos humanos e pelo

volume de faturação. Assim, afere-se que um fator comum às duas variáveis são os custos correntes.

A tabela 10 descreve o fator nº9. Este fator concentra a informação sobre o GAC e o GAO. Ambas as variáveis estão bem representadas no fator e os valores registados na coluna das comunalidades denotam uma elevada capacidade explicativa do fator. A capacidade de variância explicativa é de 2,928% (ver tabela 1) e o valor próprio é de 1,347. Tanto o GAC como GAO têm associado à sensibilidade dos resultados líquidos correntes como à volatilidade dos resultados operacionais as variações nas vendas, tornando assim em comum neste fator as variações nas vendas. Designa-se este fator como a componente da “Volatilidade dos Resultados Operacionais”.

Tabela 10: “Volatilidade dos Resultados Operacionais”

Componente Principal 9 – “Volatilidade dos Resultados Operacionais”				
Valor Próprio	1,347			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore: GAC t	,851		,725	0,809
Zscore: GAO t	,822		,676	0,818
Componente da "Volatilidade dos Resultados Operacionais"				

O fator nº 10 é designado pelo fator da “Sensibilidade dos Resultados Correntes”. Esta componente principal é responsável pela retenção da informação de variável V2 e GAF (ver tabela 11).

Tabela 11: “Sensibilidade dos Resultados Correntes”

Componente Principal 10 –“Sensibilidade dos Resultados Correntes”				
Valor Próprio	1,215			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore(V2)	,813		0,660	0,708
Zscore: GAF t		-,842	0,709	0,721
Componente da "Sensibilidade dos Resultados Correntes"				

Ambas as variáveis apresentam um índice de conteúdo de informação elevado e estão bem representadas no eixo, Apresentam um valor próprio de 1,215 e uma variância explicativa de 2,641%.

Em comum as duas variáveis deste fator têm a capacidade de libertação de meios, tanto para contrapor com o passivo como com os resultados operacionais.

O fator nº 11 é o fator que apenas extrai a variável vendas, como se pode ver na tabela 12.

Tabela 12: “Variabilidade das Vendas”

Componente Principal 11 – “Variabilidade das Vendas”				
Valor Próprio	1,049			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore: Vendas t	,711		0,505	0,622
Componente da "Variabilidade das Vendas"				

A variável “Vendas” apresenta um nível de correlação forte com o fator e está bem representada no eixo da componente principal. Este explica 1,049 de valor próprio e uma variância explicada de 2,281%. Este fator é designado por “Variabilidade das Vendas”.

O fator nº 12 mede a “Capacidade de Reembolso” das empresas e está representado na tabela 13.

Tabela 13: “Capacidade de Reembolso”

Componente Principal 12 – “Capacidade de Reembolso”				
Valor Próprio	1,038			
Variáveis	(+)	(-)	C.Individual	Comunalidade
Zscore(V4)	,886		0,785	0,823
Componente da "Capacidade de Reembolso"				

A variável V4 apresentam um nível de correlação muito forte com o fator e a informação retirada também é significativa. A variância explicada é de 2,256%, conforme é visível na tabela 1, e o seu valor próprio é de 1,038.

A qualidade do modelo foi aferida com recurso aos índices de qualidade GFI, AGFI e RMSR com os valores 0,99, 0,97 e 0,01 respetivamente (anexo 1). Concluo assim pelo muito bom ajustamento do modelo aos dados da amostra.

A análise fatorial apresenta 4 variáveis que saturam dois fatores. As variáveis VAB, EBE e Capital Próprio tanto podem ser explicadas pelo fator das “Encargos e Responsabilidades” como pela rubrica da “Solvabilidade”. A rubrica CMVMC também

está presente nos fatores 1 e 5 podendo ser explicado pelo fator nº1 “Encargos e Responsabilidades” e pelo fator dos “Proveitos Operacionais Correntes”. Foi opção considerar estas variáveis no fator em que o coeficiente de correlação é superior.

A análise fatorial permitiu compactar a informação inicial em 12 fatores:

Fator 1 - “Encargos e Responsabilidades”;

Fator 2 - “Resultados Obtidos”;

Fator 3 - “Solvabilidade”;

Fator 4 - “Liquidez”;

Fator 5 - “Proveitos Operacionais Correntes”;

Fator 6 - “Rendibilidade”;

Fator 7 - “Prazos de Pagamentos”;

Fator 8 - “Custos Operacionais Correntes”;

Fator 9 - “Volatilidade dos Resultados Operacionais”;

Fator 10 - “Sensibilidade dos Resultados Correntes”;

Fator 11 - “Variabilidade das Vendas”;

Fator 12 - “Capacidade de Reembolso”.

## 9- Modelos de Regressão

Neste capítulo será descrita a capacidade preditiva e explicativa dos fatores selecionados na análise fatorial como variáveis explicativas em três contextos diferentes.

Numa primeira fase o modelo de regressão utilizado é o logístico. Este modelo terá como variáveis dependentes, “Situação” e “Incidentes”. Trata-se de duas variáveis binárias em que a primeira retrata a situação atual da empresa (Solvente=0 ou Insolvente=1) e a segunda contabiliza a existência de incidentes registados nas empresas (Sem Incidentes=0 ou Com Incidentes=1).

Numa segunda fase utilizo o modelo de regressão linear dos mínimos quadrados ordinários para descrever o comportamento da variável Endividamento. A aplicação deste modelo prende-se com o facto da variável Endividamento ser uma variável contínua.

Em ambas as situações as variáveis explicativas são os 12 fatores obtidos anteriormente e apresentados na figura 4.

Figura 4- Regressões utilizadas neste estudo.

Modelos de Regressão	Variáveis Dependentes	Variáveis Independentes
Logístico	Situação	Fator 1
	Incidentes	Fator 2
Linear	Endividamento	Fator 3
		Fator 4
		Fator 5
		Fator 6
		Fator 7
		Fator 8
		Fator 9
		Fator 10
		Fator 11
		Fator 12

### 9.1- O modelo Logit e a variável dependente “Situação”

A variável “Situação” classifica as empresas que se encontram numa situação de solvência e de insolvência.

Pretendo com o modelo Logit analisar qual a capacidade preditiva que as variáveis independentes têm na classificação correta das empresas. Como referi as variáveis independentes são os fatores que se obtiveram na análise fatorial. Também se pretende aferir qual a influência individual que cada fator tem na classificação das empresas.

Assim, a probabilidade estimada de uma empresa entrar em insolvência é dada pelo modelo final:

$$\text{Prob}(\text{Situação}=1)=\frac{1}{1+\exp -(-4,618+0,664\text{FAC1}-0,119\text{FAC2}+0,150\text{FAC3}-0,914\text{FAC4}+0,247\text{FAC5}+0,267\text{FAC6}-0,150\text{FAC7}-0,016\text{FAC8}-0,223\text{FAC9}-0,346\text{FAC10}+0,080\text{FAC11}+0,251\text{FAC12})}$$

O modelo estimado apresenta um Pseudo- $R^2$  McFadden de 21,6% e um  $P$ -value de 0,047817 (anexo 2) o que para um nível de significância de 5% conclui-se que o modelo é globalmente significativo.

A estatística -2LL (-2 Log Likelihood) de valor 76,860 e  $P$ -value = 1 e o teste Hosmer e Lemeshow com um  $X^2=3,674$  e um  $P$ -value = 0,885 (anexo 2) levam a concluir que o modelo se ajusta aos dados para qualquer nível de significância ao não ser rejeitado a hipótese nula.

O teste aos coeficientes do Omnibus de valor  $X^2=21,179$  e  $P$ -value =0,048 (anexo 2) permite concluir que existe pelo menos uma variável independente no modelo com capacidade preditiva sobre a variável dependente “Situação”.

Em termos individuais as variáveis “Encargos e Responsabilidades” com um  $X^2_{\text{Wald}} = 15,566$  e um  $P$ -value <0,001 e os “Proveitos Operacionais Correntes” com um  $X^2_{\text{Wald}} = 4,133$  e um  $P$ -value = 0,042 (ver tabela 14), associadas aos parâmetros  $\beta_2$  e  $\beta_6$  estimados, são individualmente significativos a 1% e 5%, respetivamente.

De acordo com os dados obtidos, concluo que o coeficiente de regressão estimado  $\beta_2 = 0,6642$  significa que a probabilidade de uma empresa ser “insolvente”,

tendo em consideração os “Encargos e Responsabilidades”, é superior em 0,6642 à de outra com as mesmas características, mas num estado de “solvência”.

Tabela 14: Output do SPSS da Regressão Logística: variável dependente “Situação”.

Variável	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para Exp(B)	
							Inferior	Superior
FAC1_1	,664	,168	15,566	1	,000	1,943	1,397	2,703
FAC2_1	-,119	,178	,450	1	,502	,887	,626	1,258
FAC3_1	,150	,354	,181	1	,671	1,162	,581	2,325
FAC4_1	-,914	1,207	,573	1	,449	,401	,038	4,272
FAC5_1	,247	,121	4,133	1	,042	1,280	1,009	1,623
FAC6_1	,267	,390	,471	1	,493	1,307	,609	2,805
FAC7_1	-,150	,427	,124	1	,725	,860	,372	1,989
FAC8_1	-,016	,190	,007	1	,935	,985	,678	1,430
FAC9_1	-,223	,585	,145	1	,703	,800	,254	2,517
FAC10_1	-,346	,476	,528	1	,468	,708	,278	1,799
FAC11_1	,080	,291	,075	1	,784	1,083	,613	1,915
FAC12_1	,251	,300	,702	1	,402	1,286	,714	2,316
Constante	-4,618	,504	84,080	1	,000	,010		

Da mesma forma, o coeficiente estimado  $\beta_6=0,2466$  representa que a probabilidade de uma empresa ser “insolvente”, tendo em conta os “Proveitos Operacionais Correntes”, é superior em 24,66 pontos percentuais à de outra com as mesmas características, mas numa situação de “solvência”.

Por forma a melhor interpretar os parâmetros estimados  $\beta_2$  e  $\beta_6$  associados às variáveis “Encargos e Responsabilidades” e “Proveitos Operacionais Correntes” recorro à interpretação da exponencial desses coeficientes.

O exponencial do coeficiente  $\beta_2$  ( $\text{Exp}(\beta_2)$ ) é de 1,943 o que permite concluir, em termos percentuais, que a probabilidade de uma empresa entrar num estado insolvente é de 94,3%,  $[100 * (1,943-1)]$ , quando se verifica um acréscimo unitário nos “Encargos e Responsabilidades”.

Também para um intervalo de confiança de 95% e um nível de significância de 5% o intervalo inferior e superior do ( $\text{Exp}(\beta_2)$ ) é de ]1,397; 2,703[, respetivamente, e sendo o intervalo inferior superior a 1, conclui-se que a influência dos “Encargos e

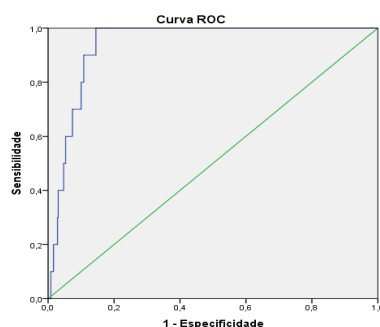
Responsabilidades” sobre a probabilidade de uma empresa entrar num estado insolvente é estatisticamente significativa.

Por outro lado,  $(\text{Exp}(\beta_6))$  é de 1,280 e em termos percentuais de 28%,  $[100 * (1,280-1)]$ , significando que a probabilidade de uma empresa entrar num estado insolvente tende a aumentar 28% quando há uma variação unitária positiva nos “Proveitos Operacionais Correntes”.

O intervalo do  $(\text{Exp}(\beta_6))$  é de  $]1,009; 1,623[$  e permite concluir que, sendo o intervalo inferior superior a 1, a influência dos “Proveitos Operacionais Correntes” sobre a probabilidade de uma empresa se tornar insolvente é estatisticamente significativa para um intervalo de confiança de 95% e um nível de significância de 5%.

O gráfico 5 apresenta a curva ROC efetuada para a variável “Situação” em função dos “Encargos e Responsabilidades”, tratando-se assim da probabilidade de uma empresa entrar em insolvência em função dos “Encargos e Responsabilidades”.

Gráfico 5: A variável “Situação” em função dos “Encargos e Responsabilidades”.



A área sob a curva ROC é de 0,939 (anexo 3) que sendo superior a 0,5 ( $P\text{-value} < 0,001$ ) permite concluir que a variável “Encargos e Responsabilidades” apresenta uma capacidade discriminante excelente.

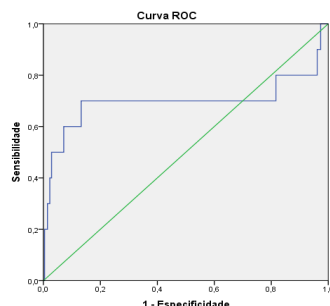
O gráfico 6 apresenta a curva ROC efetuada para a variável “Situação” em função dos “Proveitos Operacionais Correntes”, como forma de avaliar a probabilidade de uma empresa entrar em insolvência em função dos “Proveitos Operacionais Correntes”.

A área sob a curva ROC é de 0,697 (anexo 3) que sendo superior a 0,5 ( $P\text{-value} = 0,033$ ) permite concluir que a variável “Proveitos Operacionais Correntes” apresenta



uma capacidade discriminante a um nível de significância de 5%, sendo assim menor que o verificado nos “Encargos e Responsabilidades”.

Gráfico 6: A variável “Situação” em função dos “Proveitos Operacionais Correntes”



As variáveis utilizadas na regressão levam a que o modelo apresente uma taxa de erro de apenas 2,60% e uma taxa de acerto de 97,40%. Das 490 empresas solventes, o modelo classifica 3, erradamente, como sendo insolventes e as 10 empresas insolventes são classificadas todas corretamente, como se vê na tabela 15.

Tabela 15: Output com a capacidade da preditiva do modelo: variável dependente “Situação”.

Expectation-Prediction Evaluation for Binary Specification						
Equation: FATORES_LOGIT						
Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	487	10	497	490	10	500
P(Dep=1)>C	3	0	3	0	0	0
Total	490	10	500	490	10	500
Correct	487	0	487	490	0	490
% Correct	99.39	0.00	97.40	100.00	0.00	98.00
% Incorrect	0.61	100.00	2.60	0.00	100.00	2.00
Total Gain*	-0.61	0.00	-0.60			
Percent Gain**	NA	0.00	-30.00			
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	480.96	9.04	490.00	480.20	9.80	490.00
E(# of Dep=1)	9.04	0.96	10.00	9.80	0.20	10.00
Total	490.00	10.00	500.00	490.00	10.00	500.00
Correct	480.96	0.96	481.93	480.20	0.20	480.40
% Correct	98.16	9.64	96.39	98.00	2.00	96.08
% Incorrect	1.84	90.36	3.61	2.00	98.00	3.92
Total Gain*	0.16	7.64	0.31			
Percent Gain**	7.80	7.80	7.80			
*Change in "% Correct" from default (constant probability) specification						
**Percent of incorrect (default) prediction corrected by equation						

Desta forma concluo que o modelo apresenta uma sensibilidade de acerto de 99,39% e uma especificidade de classificação de 97,99%.

As mesmas variáveis independentes foram utilizadas para prever a ocorrência de incidentes nos pagamentos nas empresas.

## 9.2- O Modelo Logit e a variável dependente “Incidentes”

A tabela 16 dá-nos o output da regressão logística com a dependente “Incidentes”.

Tabela 16: Output do SPSS da Regressão Logística: variável dependente “Incidentes”

Variável	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para Exp(B)	
							Inferior	Superior
FAC1_1	,157	,122	1,659	1	,198	1,170	,921	1,486
FAC2_1	-,483	,160	9,137	1	,003	,617	,451	,844
FAC3_1	-,044	,124	,124	1	,725	,957	,751	1,220
FAC4_1	-,537	,292	3,374	1	,066	,585	,330	1,037
FAC5_1	-,243	,210	1,338	1	,247	,784	,520	1,184
FAC6_1	,082	,123	,447	1	,504	1,086	,853	1,382
FAC7_1	,620	,108	32,790	1	,000	1,859	1,503	2,298
FAC8_1	,015	,121	,016	1	,900	1,015	,802	1,286
FAC9_1	,035	,126	,078	1	,780	1,036	,809	1,326
FAC10_1	,026	,114	,051	1	,822	1,026	,820	1,284
FAC11_1	-,050	,113	,194	1	,659	,951	,762	1,188
FAC12_1	,044	,125	,122	1	,726	1,045	,818	1,334
Constante	-1,429	,124	131,815	1	,000	,240		

A regressão mostra-nos que a probabilidade estimada de uma empresa registar incidentes é dada pelo modelo:

$$\text{Prob(Incidentes=1)} = \frac{1}{1 + \exp(-(-1,429 + 0,157\text{FAC1} - 0,483\text{FAC2} - 0,044\text{FAC3} - 0,537\text{FAC4} - 0,243\text{FAC5} + 0,082\text{FAC6} + 0,620\text{FAC7} + 0,015\text{FAC8} + 0,035\text{FAC9} + 0,026\text{FAC10} - 0,050\text{FAC11} + 0,044\text{FAC12}))}$$

O modelo estimado apresenta um Pseudo- $R^2$  McFadden de 10,49% e um  $P$ -value de 0,0000 <0,001 (anexo 4), o que para qualquer nível de significância conclui que o modelo é globalmente significativo.

A estatística -2LL (-2 Log Likelihood) de valor 473,902 e  $P$ -value = 0,67 leva a concluir que o modelo se ajusta aos dados.

O teste Hosmer e Lemeshow com um  $X^2=19,662$  e um  $P$ -value = 0,012 permite verificar que os valores estimados pelo modelo não são próximos dos valores observados. Estes resultados podem ser explicados pelas variáveis explicadas terem sido extraídas pela análise fatorial tendo como variável de decisão a “Situação” das empresas.

O teste aos coeficientes do Omnibus de valor  $X^2=55,526$  e  $P$ -value <0,001 permite concluir que existe pelo menos uma variável independente no modelo com capacidade preditiva sobre a variável dependente “Situação”.

Em termos individuais as variáveis “Resultados Obtidos” com um  $X^2_{\text{Wald}}=9,137$  e  $P$ -value = 0,003, a “Liquidez” com um  $X^2_{\text{Wald}}=3,374$  e  $P$ -value = 0,066 e os “Prazos de Pagamentos” com  $X^2_{\text{Wald}}=32,790$  e um  $P$ -value <0,001, associadas aos parâmetros estimados  $\beta_3$ ,  $\beta_5$  e  $\beta_8$  são individualmente significativos para um nível de significância de 1% e 10%.

O coeficiente de regressão estimado  $\beta_3 = -0,483$  mostra que a probabilidade de uma empresa registar incidentes, tendo em consideração os “Resultados Obtidos”, é inferior em 0,4834 à de outra com as mesmas características, mas sem incidentes.

O coeficiente  $\beta_5 = -0,537$  tende a indicar que a probabilidade de uma empresa registar incidentes, tendo em conta a “Liquidez”, é inferior em 0,537 à de outra empresa com as mesmas características, mas sem incidentes.

O coeficiente  $\beta_8 = 0,620$  tende a mostrar que a probabilidade de uma empresa registar incidentes, tendo em conta o “Prazo de Pagamentos”, é superior em 62 pontos percentuais à de outra empresa com as mesmas características, mas sem incidentes.

Por forma a verificar o impacto das variáveis explicativas sobre a probabilidade da ocorrência de incidentes calcularam-se as exponenciais das estimativas para cada um dos coeficientes. Neste caso vou interpretar as exponenciais para  $\beta_3$ ,  $\beta_5$  e  $\beta_8$ .

O exponencial do coeficiente  $\beta_3$  (Exp ( $\beta_3$ )) é de 0,617 permitindo concluir que a probabilidade de uma empresa registrar incidentes diminui em 38,3%  $[100 * (0,617 - 1)]$  quando se verifica um aumento nos “Resultados Obtidos”.

Para um intervalo de confiança de 95%, o rácio das probabilidades da variável “Resultados Obtidos” [Exp ( $\beta_3$ )] situa-se no intervalo  $]0,451; 0,844[$  podendo afirmar-se que, para um nível de significância de 5%, a influência dos “Resultados Obtidos” sobre a probabilidade de ter um incidente é estatisticamente significativa.

No mesmo seguimento o Exp ( $\beta_5$ ) é de 0,585 permitindo aferir que a probabilidade de uma empresa registrar incidentes diminui em 41,5%  $[100 * (0,585 - 1)]$  quando se verifica um aumento da “Liquidez”.

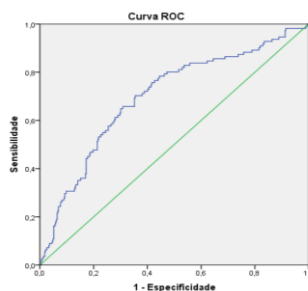
O intervalo do Exp ( $\beta_5$ ) é de  $]0,330; 1,037[$  o que para um intervalo de confiança de 95% e um nível de significância de 5% a “Liquidez” tende a não ser significativa. Este resultado era esperado uma vez que a “Liquidez” é individualmente significativos para um intervalo de confiança a 93%, veja-se o  $P\text{-value} = 0,066$ .

Por último, o Exp ( $\beta_8$ ) é de 1,859 o que significa que a probabilidade de uma empresa registrar incidentes tende a aumentar em 85,9%  $[100 * (1,859 - 1)]$  quando se verifica um aumento nos “Prazos de Pagamentos”.

O intervalo do Exp ( $\beta_8$ ) é  $]1,503; 2,298[$  concluindo-se que para o intervalo de confiança de 95% e um nível de significância de 5%, a influência dos “Prazos de Pagamentos” na probabilidade da ocorrência de incidentes é estatisticamente significativa.

O Gráfico 7 apresenta uma curva ROC sobre a variável “Incidentes” em função da variável “Prazos de Pagamentos”. Esta curva é utilizada com o intuito de se avaliar a capacidade discriminante de uma empresa na probabilidade de esta registrar incidentes tendo em conta os “Prazos de Pagamentos”.

Gráfico 7: A variável “Incidentes” em função dos “Prazos de Pagamentos”.



A área da curva ROC é de 0,698 (anexo 4) que sendo superior a 0,5 e para um *P-value* <0,001 então a variável “Prazos de Pagamentos” apresenta capacidade discriminante para qualquer nível de significância.

O modelo apresenta uma taxa de acerto de 77,80% e uma taxa de erro de 22,20%, como se verifica na tabela 17.

Tabela 17: Output da predição do modelo: variável dependente “Incidentes”.

Expectation-Prediction Evaluation for Binary Specification						
Equation: INCIDENTES_LOGIT						
Success cutoff: C = 0.5						
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
P(Dep=1)≤C	372	94	466	389	111	500
P(Dep=1)>C	17	17	34	0	0	0
Total	389	111	500	389	111	500
Correct	372	17	389	389	0	389
% Correct	95.63	15.32	77.80	100.00	0.00	77.80
% Incorrect	4.37	84.68	22.20	0.00	100.00	22.20
Total Gain*	-4.37	15.32	0.00			
Percent Gain**	NA	15.32	0.00			
	Estimated Equation			Constant Probability		
	Dep=0	Dep=1	Total	Dep=0	Dep=1	Total
E(# of Dep=0)	312.58	76.42	389.00	302.64	86.36	389.00
E(# of Dep=1)	76.42	34.58	111.00	86.36	24.64	111.00
Total	389.00	111.00	500.00	389.00	111.00	500.00
Correct	312.58	34.58	347.15	302.64	24.64	327.28
% Correct	80.35	31.15	69.43	77.80	22.20	65.46
% Incorrect	19.65	68.85	30.57	22.20	77.80	34.54
Total Gain*	2.55	8.95	3.97			
Percent Gain**	11.50	11.50	11.50			
*Change in "% Correct" from default (constant probability) specification						
**Percent of incorrect (default) prediction corrected by equation						

Das 389 empresas sem registo de incidentes o modelo classifica 17 como tendo incidentes e das 111 que apresentam incidentes o modelo classifica 17 como não tendo incumprimentos. Desta forma a sensibilidade, medida pela taxa de positivos verdadeiros, é de 4,37% e a especificidade, obtida pela taxa de negativos verdadeiros, é de 79,83%. A taxa de fracasso é de 50% e a taxa de falso alarme é de 20,17%. O modelo apresenta ainda um valor preditivo positivo de 15,32% e um valor preditivo negativo de 95,63%.

### 9.3- Modelo OLS e a variável dependente “Endividamento”

Nesta secção pretendo verificar como se comportam as variáveis explicativas quando se muda para uma variável quantitativa, neste caso “Endividamento”.

A estimação dos coeficientes do modelo especificado foi efetuado via Mínimos Quadrados de Ordinários com recurso ao estimador consistente White. Justifica-se pelo motivo da existência de heteroscedasticidade detetada pela execução do teste de White e de Breush-Pagan-Godfrey em que o Qui-quadrado observado é sempre superior ao Qui-quadrado crítico para qualquer nível de significância.

O output do modelo estimado encontra-se na tabela 18 e é dado por:

$$\text{Incidentes}_i = 82,260 - 3,793\text{FAC1}_i - 4,685\text{FAC2}_i - 22,235\text{FAC3}_i - 3,340\text{FAC4}_i - 1,382\text{FAC5}_i - 3,496\text{FAC6}_i + 1,579\text{FAC7}_i - 1,141\text{FAC8}_i + 0,498\text{FAC9}_i + 0,710\text{FAC10}_i - 0,209\text{FAC11}_i + 0,880\text{FAC12}_i$$

As variáveis independentes são os fatores obtidos na análise fatorial.

Com um  $R^2 = 0,904212$ , um  $F_{\text{obs}} = 383,0937$  e um  $p\text{-value} = 0,00000$  concluo, para um nível de significância de 1%, a significância global da regressão.

Para um nível de significância de 1% e face à informação estatística disponível, as variáveis “Encargos e Responsabilidades”, “Resultados Obtidos”, “Solvabilidade”, “Liquidez”, “Rendibilidade”, “Prazos de Pagamentos”, “Sensibilidade dos Resultados Correntes” e “Capacidade de Reembolso” são individualmente significativas.

Para um nível de significância de 5%, as variáveis “Proveitos Operacionais Correntes”, “Custos Operacionais Correntes” e a “Volatilidade dos Resultados Operacionais” são também individualmente significativos.

Uma variação positiva nas variáveis “Encargos e Responsabilidades”, e Resultados Obtidos”, “Solvabilidade”, “Liquidez”, “Proveitos Operacionais Correntes”, “Rendibilidade” e os “Custos Operacionais Correntes” tende a diminuir o nível de endividamento. Inversamente uma variação positiva nas variáveis “Prazos de Pagamento”, “Volatilidade dos Resultados Operacionais”, “Sensibilidade dos Resultados Correntes” e a “Capacidade de Reembolso” tende a aumentar o nível de endividamento das empresas.

Das 12 variáveis incluídas no modelo, as variáveis associadas aos parâmetros estimados  $\beta_4$  “Solvabilidade” e  $\beta_8$  “Prazos de Pagamentos” são as que mais contribuem para a diminuição e aumento do nível de endividamento. Uma variação unitária positiva na “Solvabilidade” tende a contribuir para uma diminuição de 22,24% do nível de endividamento, *Ceteris Paribus*. Da mesma forma, por cada dia adicional no atraso a liquidação dos pagamentos faz com que o nível de endividamento tenda a aumentar em 1,58%, mantendo tudo o resto constante.

Tabela 18: Output do Eviews regressão OLS: variável dependente “Endividamento”

Dependent Variable: ENDVIDA				
Method: Least Squares				
Date: 07/31/14 Time: 22:55				
Sample: 1 500				
Included observations: 500				
White heteroskedasticity-consistent standard errors & covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	82.26038	0.349106	235.6315	0.0000
FAC1_1	-3.793190	0.785202	-4.830848	0.0000
FAC2_1	-4.684514	0.517209	-9.057297	0.0000
FAC3_1	-22.23535	1.464769	-15.18011	0.0000
FAC4_1	-3.339803	0.910450	-3.668298	0.0003
FAC5_1	-1.382063	0.556076	-2.485384	0.0133
FAC6_1	-3.496093	0.555306	-6.295800	0.0000
FAC7_1	1.579481	0.564318	2.798920	0.0053
FAC8_1	-1.141008	0.531969	-2.144875	0.0325
FAC9_1	0.498039	0.222241	2.240991	0.0255
FAC10_1	0.709996	0.228471	3.107591	0.0020
FAC11_1	-0.208804	0.645227	-0.323613	0.7464
FAC12_1	0.879861	0.293658	2.996210	0.0029
R-squared	0.904212	Mean dependent var	82.26038	
Adjusted R-squared	0.901851	S.D. dependent var	24.91726	
S.E. of regression	7.806248	Akaike info criterion	6.973382	
Sum squared resid	29676.57	Schwarz criterion	7.082962	
Log likelihood	-1730.346	Hannan-Quinn criter.	7.016381	
F-statistic	383.0937	Durbin-Watson stat	1.940388	
Prob(F-statistic)	0.000000			

## **10- Interpretação dos resultados**

A análise fatorial permitiu reduzir o número de 44 variáveis iniciais em apenas 12 fatores. Esta técnica permite organizar a informação inicial em menos fatores e assim possibilitar uma melhor leitura da mesma. Os testes de qualidade do ajustamento permitem concluir que existe um ótimo ajustamento dos fatores com a informação da amostra inicial. A informação contida nos 12 fatores representa 80% de toda a informação original.

O primeiro fator “Encargos e Responsabilidades” é o resultado da junção de 15 variáveis e responsável por 36,10% de toda a variância explicada. Na sua maioria as variáveis apresentam um grau de correlação muito forte. Os Resultados Financeiros contrapõem-se, entre outras, ao Passivo, Juros suportados, à contabilização das dívidas de curto prazo e de terceiros bem como a dimensão do número de funcionários com o volume de faturação. Dois aspetos a reter: um prende-se à boa representatividade das variáveis no fator; outro à extração da informação original que é muito fidedigna à informação inicial. O fator “Encargos e Responsabilidades” é o que mais informação retém e o que melhor prediz corretamente a classificação da empresa, solvente ou insolvente, quando usado como variável independente na regressão logística na probabilidade de ocorrência de insolvência. Este fator apresenta um sinal positivo quando utilizado na predição da ocorrência de incidentes o que não contradiz a evidência empírica. Quanto maior os “Encargos e Responsabilidades” maior a propensão para ocorrência de incidentes. Por fim, quando aos “Encargos e Responsabilidades” são usados como variável no estudo da variabilidade do nível de endividamento, têm um peso negativo o que também vai de encontro ao esperado. Sendo que em média as empresas que compõem a amostra apresentam um nível de endividamento superior a 80%, conclui-se que à medida que os “Encargos e Responsabilidades” aumentam, as empresas deixam de ter margem para se poderem endividar e de certo modo tenderão a diminuir a exposição ao financiamento já existente.

O fator 2 mede os “Resultados obtidos” e é representativo da informação de 8 variáveis cuja informação original foi extraída a um nível superior a 50% e a sua representatividade está muito bem presente no fator. As variáveis apresentam um grau



de correlação muito forte e a variável que apresenta um maior “peso” é o “Resultado Líquido”. Ainda assim, é também relevante ver o contributo dos resultados correntes e antes de impostos. O nível de lucro contabilizado por cada empregado e os saldos médios dos depósitos bancários e caixa deve ser visto como indicadores de rentabilidade e liquidez imediata. Na interpretação como fator deve ser retida a capacidade de explicar 10% de toda a inércia. Como variável explicativa, apresenta sinais negativos na interpretação da variável “Situação”, “Incidente” e “Endividamento”. Não deixa de contradizer a evidência empírica porque em todos os cenários os níveis de “Resultados” positivos tenderam a contribuir para a solvabilidade de uma entidade empresarial, a diminuição da possibilidade de incumprimentos e do próprio nível de endividamento. Este último aspeto poderia ser diferente se as empresas denotassem um nível de endividamento inferior à média existente. Resultados positivos, em conjugação com a capacidade da entidade libertar meios, pode levar ao aumento do endividamento, fruto da capacidade de poder obter crédito.

O terceiro fator representa a “Solvabilidade” da empresa. Opõe a Autonomia Financeira ao nível de Endividamento. É um fator com uma capacidade explicativa de 6% e as duas variáveis que o compõem estão bem representadas. Esta contribui positivamente na classificação do estado da empresa mas apresenta um sinal negativo na contabilização de possíveis ocorrências de incidentes e de aumento de exposição ao financiamento externo. Conclui-se que os resultados obtidos nesta variável se adequam à evidência empírica. Uma entidade que apresenta capitais próprios capazes de atenuar a exposição ao endividamento leva à não possível ocorrência de incidentes e à diminuição do financiamento alheio.

O fator 4, “Liquidez”, mede o nível de liquidez reduzida e imediata que a empresa tem em determinado período. Este fator explica 4,22% da variância e as duas variáveis estão bem representadas no eixo. Quando usado como variável independente, este apresenta sinais negativos, tanto ao nível da “Situação”, do “Incidente” como do “Endividamento”. O nível de liquidez deve ser visto como um indicador da empresa em gerar meios líquidos e representa a existência de fundos para liquidar responsabilidades de curto prazo.

O fator 5, “Proveitos Operacionais Correntes” é formado pela variável “Proveitos Operacionais por Empregado” e pela variável “Custo da Mercadoria

Vendida e Matéria Consumida”. Este fator extrai 93% de toda a informação inicial de cada variável indicando ser portador de informação das variáveis muito completa. Ao ser transformado em variável explicativa, os “Proveitos Operacionais Correntes” contribuem positivamente na probabilidade de ocorrência de insolvência e negativamente nos incidentes e variação do nível de endividamento. Esta variável contabiliza os proveitos produtivos ao nível da mão-de-obra e o do custo da matéria-prima usada no processo de fabrico.

O fator 6 define-se como “Rendibilidade”. Este mede a rendibilidade do ativo, do capital próprio e financeira. Contribui positivamente na probabilidade de classificar corretamente uma empresa, na probabilidade da existência de incidente e negativamente na variabilidade do nível de endividamento. A rendibilidade mede a capacidade da empresa rentabilizar o ativo, capital próprio e os investimentos financeiros. Tendo como ponto de partida esta interpretação conclui-se que o sinal associado ao “peso” desta variável em cada uma das três situações é o mais assertivo.

O fator 7 representa os “Prazos de Pagamentos”, não só em termos de pagamentos mas também ao nível de recebimento. Quando utilizado como variável explicativa no modelo de regressão logística, tendo como variável dependente “Situação”, o sinal é negativo. Significa que o nível da probabilidade da existência de insolvência diminui. Inversamente, quanto maior for esta variável maior é a probabilidade da ocorrência de incidentes. Ao nível da variação do “Endividamento” os “Prazos de Pagamentos” denotam que quanto maior o número de dias de recebimento e de pagamento mais elevado é o nível de “Endividamento”. Estes sinais correspondem aos esperados. Um aumento do prazo médio de recebimento leva a uma gestão mais apertada da tesouraria.

Designa-se o fator 8 como os “Custos Operacionais Correntes”. Este fator mede o peso do Passivo no Volume de Negócios, variável V1, e também o custo que a empresa tem com a mão-de-obra ao ano. Este fator, quando utilizado como variável explicativa, apresenta sinais negativos na predição do estado da empresa e na variabilidade no nível de endividamento. Inversamente apresenta um sinal positivo na probabilidade de ocorrência de incidentes. Os sinais tendem a corresponder ao esperado na medida em que um elevado passivo perante o volume de negócios e custos elevados ao nível da mão-de-obra levam uma empresa a estar mais propensa à ocorrência de

insolvência, à ocorrência de incidentes e à diminuição da exposição ao financiamento por não ter condições para o efeito.

As variáveis presentes no fator 9, GAC e GAO, levam a designar este como a “Volatilidade dos Resultados Operacionais”. Este fator tem uma capacidade explicativa de 2,9% e a informação retida nas variáveis é expressivo. Esta informação deve ser vista como a importância que os resultados operacionais têm nas empresas. Quando usado como variável independente apresenta um sinal negativo na interpretação da “Situação” e sinal positivo na explicação da ocorrência de “Incidentes” e “Endividamento”. À medida que os resultados operacionais e resultados líquidos aumentam face à variabilidade das vendas, o sinal do “peso” que este fator toma na explicação da “Situação” e no “Endividamento” evidencia corretamente o esperado. O mesmo não ocorre na interpretação dos “Incidentes” onde o sinal positivo é contrário ao esperado pois o aumento dos resultados obtidos tenderam a fortalecer a empresa de meios para não ocorrer incidentes. Esta situação não se evidencia no resultado final obtido.

O fator 10 é definido como a “Sensibilidade dos Resultados Correntes”. Este contabiliza o peso que o Passivo tem nos resultados antes de impostos e a sensibilidade dos resultados líquidos perante os resultados operacionais. A sinalização negativa na probabilidade da ocorrência de incidentes tenderá a ser contrária ao esperado. Ainda assim uma possível explicação pode ser a dimensão que o passivo assume nos resultados antes de impostos. O sinal negativo na probabilidade de uma empresa ser insolvente e o sinal positivo do coeficiente no nível do endividamento correspondem ao esperado.

O fator 11 é constituído apenas pela variável “vendas” e designa-se pela “Variabilidade das Vendas”. Como variável explicativa esta apresenta um sinal positivo na interpretação da “Situação” e do “Endividamento”. Quanto à explicação dos “Incidentes” o sinal é negativo. Sendo a variabilidade das vendas um espelho da performance da atividade da empresa o sinal que apresenta nos vários cenários interpretativos corresponde ao esperado.

O fator 12 mede a “Capacidade de Reembolso” das empresas. Este fator mede o diferencial do passivo ao ativo e o peso desta diferença nos resultados líquidos e amortizações. Esta variável é responsável por 2,26% da variância explicada e apresenta um contributo individual superior a 50%. Quando utilizada como variável explicativa a

“Capacidade de Reembolso” responde positivamente à interpretação da variável “Situação”, “Incidentes” e “Endividamento”. Em termos interpretativos o sinal presente no peso da probabilidade de ocorrência de “Incidentes” é positivo mas seria de esperar um sinal negativo. Uma explicação poderá estar relacionada com a dimensão do passivo em relação ao ativo e aos próprios resultados esperados.

Os resultados extraídos na análise fatorial e a sua aplicabilidade levam a concluir que este método é útil na reorganização da informação financeira.

## 11- Conclusão

A análise fatorial correspondeu às expectativas iniciais. A sua aplicabilidade permitiu agrupar as 44 variáveis iniciais em apenas 12 fatores sem perda significativa da informação original que envolve cada variável. A seletividade das variáveis para cada fator foi calculada pelo nível de correlação e o seu contributo individual é significativo.

O uso dos fatores como variáveis explicativas corresponde corretamente aos pressupostos. Ao nível da probabilidade de uma empresa se tornar insolvente apenas as variáveis “Encargos e Responsabilidades” e “Proveitos Operacionais Correntes” são individualmente significativas e o modelo Logit apenas classifica erradamente 3 casos em 490 possíveis. Ao nível interpretativo da probabilidade de ocorrência de “Incidentes” o modelo Logit apresenta uma taxa de acerto de 77,80%, o que corresponde a uma classificação errada de 17 empresas. No que respeita à evolução do nível de endividamento a inferência estatística é feita assintoticamente e das 12 variáveis incluídas no modelo linear dos mínimos quadrados apenas a variável “Variabilidade das Vendas” não é individualmente significativa para um nível de significância a 5%.

De todos os fatores calculados, o de “Encargos e Responsabilidades” é o que apresenta uma capacidade de inércia explicada mais elevada.

Ao nível da limitação do modelo, a amostra de empresas consideradas no estudo, apesar de ser estatisticamente representativa, um número superior de empresas poderá levar a interpretações mais específicas, nomeadamente na seleção de fatores.

De futuro, será interessante aplicar a mesma análise fatorial aos dois grupos de empresas, solventes e insolventes, e verificar o comportamento dos fatores nas duas realidades.

Outras técnicas, como a análise discriminante pode ser utilizada para analisar os indicadores financeiros e prever os sinais de falência. Na presente amostra a análise discriminante não pode ser aplicada devido ao reduzido número de empresas em situação “Insolvente”.

## 12- Referências bibliográficas

- Altman, E.I., 1968. Financial ratios discriminante analysis and the prediction of corporate bankruptcy. *Journal of Finance* 589-609.
- Altman, E.I., Haldeman, R., Narayanan, P., 1977. Zeta analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* 29–54.
- Altman, E.I., Saunders, A., 1997. Credit risk measurement: Developments over the last twenty years. *Journal of Banking and Finance* 21, 1721–1742.
- Batista, A. (2012). *Credit Scoring – Uma Ferramenta de Gestão Financeira* (1ª edição), Porto: Vida Económica.
- Beaver, W., 1967, Financial ratios as predictors of failure, Empirical research in accounting: Selected studies, 1966, *Journal of Accounting Research*, V, supplement.
- Crouhy, M., Mark, R., (2000). “A comparative analysis of current credit risk models.” *Journal of Banking and Finance* 24 (1/2), 59–117, this issue.
- Fernandes, C., Peguinho, C., Vieira, E. e Neiva, J. (2013). “Análise Financeira – Teoria e Prática”. (2ª edição), Lisboa: Edições Sílabo.
- Gordy, M.B., (2000). “A comparative anatomy of credit risk models.” *Journal of Banking and Finance* 24 (1/2), 119–149, this issue.
- Grunert, K. G., et al. (2004). "Consumer perception of meat quality and implications for product development in the meat sector - a review." *Meat Science* 66(2): 259-272.
- Johnston, J. e al. (2001). *Métodos Econométricos* (4ª Edição), Lisboa: McGraw-Hill.
- Lawrence, E.L., Smith, S., Rhoades, M., (1992). “An analysis of default risk in mobile home credit.” *Journal of Banking and Finance* 299–312.
- Lehmann, B., “Is it worth the while? The relevance of qualitative information in credit rating”, Working Paper presented at the EFMA 2003 Meetings, Helsinki, 2003.
- Lopez, J. A. and M. R. Saidenberg (2000). “*Evaluating credit risk models.*” *Journal of Banking & Finance* 24(1-2): 151-165.
- Martin, D., (1977). “Early warning of bank failure: A logit regression approach.” *Journal of Banking and Finance* 249–276.

Moreira, J.(2001), *Análise Financeira de Empresas: da Teoria à Prática*, 4ª edição, Porto: Instituto do Mercado de Capitais.

Maroco, J.(2011), *Análise de Estatística com o SPSS Statistics*, 5ª edição ReportNumber,Lda. Pero Pinheiro.

Oliveira, M. e al. (2011). *Econometria* (1ª Edição),Lisboa: Escolar Editora.

Platt, H.D., Platt, M.B., (1991). “A note on the use of industry-relative ratios in bankruptcy prediction.” *Journal of Banking and Finance* 1183–1194.

Smith, L.D., Lawrence, E., (1995). “Forecasting losses on a liquidating long-term loan portfolio.” *Journal of Banking and Finance* 959–985.

Reis, E. (2001). *Estatística Multivariada Aplicada* (2ª Edição), Lisboa: Edições Sílabo.

West, R.C., (1985). “A factor-analytic approach to bank condition.” *Journal of Banking and Finance* 253–266.

[www.bportugal.pt](http://www.bportugal.pt), acedido em Janeiro de 2014.

[www.ine.pt](http://www.ine.pt), acedido em Janeiro de 2014.

[www.dn.pt](http://www.dn.pt), acedido em Dezembro de 2013.

[www.tvi.pt](http://www.tvi.pt), acedido em Dezembro de 2013.

### 13- Anexos

#### Anexo 1: Matriz das Correlações

		Zscore: VN t	Zscore: VAB t	Zscore: Cpessoal t	(...)	Zscore(V2)	Zscore(V3)	Zscore(V4)
Correlação reproduzida	Zscore: VN t	,923 <sup>a</sup>	,804	,687	(...)	-,018	,233	-,028
	Zscore: VAB t	,804	,960 <sup>a</sup>	,857	(...)	-,031	,358	-,027
	Zscore: Cpessoal t	,687	,857	,866 <sup>a</sup>	(...)	-,052	,354	-,008
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	Zscore(V2)	-,018	-,031	-,052	(...)	,708 <sup>a</sup>	-,087	,113
	Zscore(V3)	,233	,358	,354	(...)	-,087	,907 <sup>a</sup>	-,058
	Zscore(V4)	-,028	-,027	-,008	(...)	,113	-,058	,823 <sup>a</sup>
Resíduos <sup>b</sup>	Zscore: VN t		,004	,021	(...)	-,003	-,002	,001
	Zscore: VAB t	,004		,068	(...)	,000	-,023	-,014
	Zscore: Cpessoal t	,021	,068		(...)	,001	-,049	-,032
	(...)	(...)	(...)	(...)	(...)	(...)	(...)	(...)
	Zscore(V2)	-,003	,000	,001	(...)		,003	-,020
	Zscore(V3)	-,002	-,023	-,049	(...)	,003		-,003
	Zscore(V4)	,001	-,014	-,032	(...)	-,020	-,003	
Método de extração: análise do componente principal.								
a. Comunalidades reproduzidas								
b. Os resíduos são computados entre as correlações observadas e reproduzidas. Há 98 (9,0%) de resíduos não redundantes com valores absolutos maiores do que 0,05.								

Nota: a matriz de correlações encontra-se abreviada pelo motivo de ser demasiado grande e serve para mostrar de onde foram retirados os dados para cálculos dos índices de ajustamento.

#### Cálculo dos índices da avaliação da qualidade do modelo:

**GFI**=1-0,5\* (0,004+0,068-0,019-0,013-0,012-0,017-0,030+0,008+0,002+0,027-0,006+0,034+0,007-0,118+0,037-0,205+0,017-0,006-0,009-0,026+0,101+0,020-0,005+0,032+0,004+0,086+0,035-0,075-0,036+0,039+0,023+0,002-0,014-0,092+0,051+0,021-0,043+0,009+0,065-0,023+0,104+0,047+0,003+0,003-0,003)<sup>2</sup>=**0,9876755**

**AGFI**=1- [(46\*(46+1)/2)/((46-13)<sup>2</sup>-46-13)/2]\*(1-0,9876755)=**0,97413**

**RMSR**= ((0,004)<sup>2</sup>+(0,068)<sup>2</sup>+(-0,019)<sup>2</sup>+(-0,013)<sup>2</sup>+(-0,012)<sup>2</sup>+(-0,017)<sup>2</sup>+(-0,030)<sup>2</sup>+(0,008)<sup>2</sup>+(0,002)<sup>2</sup>+(0,027)<sup>2</sup>+(-0,006)<sup>2</sup>+(0,034)<sup>2</sup>+(0,007)<sup>2</sup>+(-0,118)<sup>2</sup>+(0,037)<sup>2</sup>+(-0,205)<sup>2</sup>+(0,017)<sup>2</sup>+(-0,006)<sup>2</sup>+(-0,009)<sup>2</sup>+(-0,026)<sup>2</sup>+(0,101)<sup>2</sup>+(0,020)<sup>2</sup>+(-0,005)<sup>2</sup>+(0,032)<sup>2</sup>+(0,004)<sup>2</sup>+(0,086)<sup>2</sup>+(0,035)<sup>2</sup>+(-0,075)<sup>2</sup>+(-0,036)<sup>2</sup>+(0,039)<sup>2</sup>+(0,023)<sup>2</sup>+(0,002)<sup>2</sup>+(-0,014)<sup>2</sup>+(-0,092)<sup>2</sup>+(0,051)<sup>2</sup>+(0,021)<sup>2</sup>+(-0,043)<sup>2</sup>+(0,009)<sup>2</sup>+(0,065)<sup>2</sup>+(-0,023)<sup>2</sup>+(0,104)<sup>2</sup>+(0,047)<sup>2</sup>+(0,003)<sup>2</sup>+(0,003)<sup>2</sup>+(-0,003)<sup>2</sup>/(46\*(46-1)/2))<sup>(1/2)</sup>=**0,0111**



## Anexo 2: Teste de significância e de qualidade de ajustamento

**Testes de coeficientes de modelo Omnibus**

		Qui- quadrado	df	Sig.
Etapa 1	Etapa	21,179	12	,048
	Bloco	21,179	12	,048
	Modelo	21,179	12	,048

**Resumo do modelo**

Etapa	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	76,860 <sup>a</sup>	,041	,233
a. Estimação encerrada no número de iteração 8 porque as estimativas de parâmetro mudaram em menos de ,001.			

$$\text{Pseudo } R^2 \text{ McFadden} = 21,179 / (21,179 + 76,860) = 0,2160 * 100 = 21,60\%$$

**Teste de Hosmer e Lemeshow**

Etapa	Qui- quadrado	df	Sig.
1	3,674	8	,885

### Anexo 3: Área sob a curva ROC

Variável de resultado do teste: <b>“Encargos e Responsabilidades”</b>				
Área	Modelo padrão <sup>a</sup>	Sig. assintótico	Intervalo de confiança assintótico 95%	
			Limite inferior	Limite superior
,939	,016	,000	,908	,970
a. Sob a suposição não paramétrica				
b. Hipótese nula: área real = 0,5				

Variável de resultado do teste: <b>“Proveitos Operacionais Correntes”</b>				
Área	Modelo padrão <sup>a</sup>	Sig. assintótico	Intervalo de confiança assintótico 95%	
			Limite inferior	Limite superior
,697	,128	,033	,446	,949
a. Sob a suposição não paramétrica				
b. Hipótese nula: área real = 0,5				

#### Anexo 4: Teste de significância e de qualidade de ajustamento

**Testes de coeficientes de modelo Omnibus**

		Qui- quadrado	df	Sig.
Etapa 1	Etapa	55,526	12	,000
	Bloco	55,526	12	,000
	Modelo	55,526	12	,000

**Resumo do modelo**

Etapa	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	473,902 <sup>a</sup>	,105	,161

a. Estimação encerrada no número de iteração 5 porque as estimativas de parâmetro mudaram em menos de ,001.

$$\text{Pseudo R2 McFadden} = 55,526 / (55,526 + 473,902) = 0,1049 * 100 = 10,49\%$$

**Teste de Hosmer e Lemeshow**

Etapa	Qui- quadrado	df	Sig.
1	19,662	8	,012

#### Área sob a curva ROC

Variável de resultado de teste: "Prazos de Pagamentos"				
Área	Modelo padrão <sup>a</sup>	Sig. assimptótico	Intervalo de confiança assintótico 95%	
			Limite inferior	Limite superior
,698	,029	,000	,642	,755
a. Sob a suposição não paramétrica				
b. Hipótese nula: área real = 0,5				